Weak approximation of stochastic partial differential equations: the non linear case

Arnaud Debussche

WEAK APPROXIMATION OF STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS: THE NON LINEAR CASE

ARNAUD DEBUSSCHE

Abstract. We study the error of the Euler scheme applied to a stochastic partial differential equation. We prove that as it is often the case, the weak order of convergence is twice the strong order. A key ingredient in our proof is Malliavin calculus which enables us to get rid of the irregular terms of the error. We apply our method to the case a semilinear stochastic heat equation driven by a space-time white noise.

April 8, 2008

1. Introduction

When one considers a numerical scheme for a stochastic equation, two types of errors can be considered. The strong error measures the pathwise approximation of the true solution by a numerical one. This problem has been extensively studied in finite dimension for stochastic differential equations (see for instance [20], [26], [27], [32]) and also more recently in infinite dimension for various types of stochastic partial differential equations (SPDEs) (see among others [1], [4], [6], [10], [11], [12], [13], [14], [15], [16], [17], [18], [22], [23], [29], [30], [34], [35], [36]). Another way to measure the error is the so-called weak order of convergence of a numerical scheme which is concerned with the approximation of the law of the solution at a fixed time. In many applications, this error is more relevant. Pioneering work by Milstein ([24], [25]) and Talay ([33]) have been followed by many articles (see references in the books cited above). Very few works exist in the literature for the weak approximation of solution of SPDEs. A delayed stochastic differential equation has been studied in [3]. Weak order for a SPDE has been studied only recently in [7], [8], [19]. In order to explain the novelty of the present article, let us focus on a specific example.

We consider a stochastic nonlinear heat equation in a bounded interval $I = (a, b) \subset R$ with Dirichlet boundary conditions and driven by a space-time white noise:

$$
\begin{cases} \dfrac{\partial X}{\partial t} = X_{\xi\xi} + f(X) + \sigma(X)\dot{\eta},\ \xi \in I,\ t > 0, \\ X(a,t) = X(b,t) = 0,\ t > 0, \\ X(\xi, 0) = x(\xi),\ \xi \in I. \end{cases} \tag{1.1}
$$

1991 *Mathematics Subject Classification.* 35A40, 60H15, 60H35.

Key words and phrases. Weak order, stochastic heat equation, Euler scheme.

ENS de Cachan, Antenne de Bretagne, Campus de Ker Lann, Av. R. Schuman, 35170 BRUZ, FRANCE (arnaud.debussche@bretagne.ens-cachan.fr).

Acknowledgments: Part of this work was done while the author visited the Institut Mittag-Leffler (Djursholm, Sweden) during the semester "Stochastic Partial Differential Equations".

Where f and σ are smooth Lipschitz functions from $\mathbb{R}$ to $\mathbb{R}$.

We introduce the classical abstract framework extensively used in the book [5]. We set $H = L^2(I)$, $A = \partial_{\xi\xi}$, $D(A) = H^2(I) \cap H_0^1(I)$, W is a cylindrical Wiener process so that the space-time white noise is mathematically represented as the time derivative of W. We set $f(x)(\xi) = f(x(\xi))$, $x \in H$ and define $\sigma : H \to \mathcal{L}(H)$ by $\sigma(x)h(\xi) = \sigma(x(\xi))h(\xi)$, $x, h \in H$. We then rewrite (1.1) as

$$\left\{ \begin{array}{l} dX = (AX + f(X))dt + \sigma(X)dW, \\ X(0) = x. \end{array} \right. \tag{1.2}$$

It is well known that this equation has a unique solution. We investigate the error committed when approximating this solution by the solution of the Euler scheme

$$\left\{ \begin{array}{l} X_{k+1} - X_k = \Delta t \left(AX_{k+1} + f(X_k)\right) + \sigma(X_k)\left(W((k+1)\Delta t) - W(k\Delta t)\right), \\ X_0 = x, \end{array} \right. \tag{1.3}$$

where $\Delta t = T/N$, $N \in \mathbb{N}$, $T > 0$.

The study of the weak error aims to prove bounds of the type:

$$|\mathbb{E}(\varphi(X(n\Delta t))) - \mathbb{E}(\varphi(X_n)| \leq c\Delta t^{\delta},$$

with a constant c which may depend on φ, x, N and on the various parameter in the equation. Also φ is assumed to be a smooth function on H. If such a bound is true, we say that the scheme has weak order δ. In comparison, the strong error is given by $\mathbb{E}(|(X(n\Delta t)) - X_n|)$ or $\mathbb{E}(\sup_{n=0,\dots,N} |(X(n\Delta t)) - X_n|)$. Clearly, if the scheme has strong order $\tilde{\delta}$ then it has weak order $\delta \geq \tilde{\delta}$. Indeed, the test functions φ are Lipschitz. In general, it is expected that the weak order is larger than the the strong order.

In the case of the Euler scheme applied to a stochastic differential equation, it is well known that the strong order is $1/2$ whereas the weak order is 1 (see [32]). The classical proof of this uses the Kolmogorov equation associated to the stochastic equation. The main difficulty to generalize this proof to the infinite dimensional equation (1.2) is that this Kolmogorov equation is then a partial differential equation with an infinite number of variables and involving unbounded operators (see (3.6) below). The delayed stochastic differential equation studied in [3] is an infinite dimensional problem but since the equation does not contain differential operators the Kolmogorov equation is simpler to study. In [19], a SPDE similar to (1.2) is considered but very particular test functions φ are used. They are allowed to depend only on finite dimensional projections of the unknown and the bound of the weak error involves a constant which strongly depends on the dimension. In [7], [8], the Kolmogorov equation is not used directly. A change of variable is used in order to simplify it. In [7], the stochastic nonlinear Schrödinger equation is considered and the fact that the linear Schrödinger equation generates an invertible group is used in an essential way. This is obviously wrong for the heat equation considered here. The same change of unknown works in the case of a linear equation with additive noise as shown in [8] but there it is used that the solution can be written down explicitly. We have not been able to generalize this idea to the non linear equation considered here.

We use in fact the original method developed by Talay in the finite dimensional case. The weak error is decomposed thanks to the Kolmogorov equations on each time step. Each term represents the error between the solution of the Kolmogorov equation on one time step and the approximation given by the numerical solution. Due to the presence of unbounded operators, this apparently requires a lot of smoothness on the numerical solution. The main idea here is to observe that the non smooth part of the solutions of (1.2) and (1.3) are contained in a stochastic integral. We get rid of this stochastic integral thanks to Malliavin calculus and an integration by part. We are thus able to prove that as expected the weak order is twice the strong order without artificial assumption except from a technical one on σ. We restrict our presentation to the abstract equation above, a nonlinear heat equation driven by a space-time white noise. However, our method is general and can be used for more general equations as will be shown in future articles. Also, we only consider a semi-discretization in time. A full discretization will treated in forthcoming works.

Note that the method developed here does allow to recover the result of [8]. Indeed, in the Euler scheme (1.3), the linear term is fully implicit and we cannot consider a scheme where it is partially implicit such as the theta-scheme considered in [8]. Note also that the proof below are much more complicated than in [8] and [7].

Malliavin calculus has already been used for the numerical analysis of stochastic equations. In [2], it is used to prove an expansion of the error of the Euler scheme for a stochastic differential equation under minimal assumptions on the test functions φ. This is a completely different idea and the Malliavin calculus is used completely differently. It is not clear that such ideas could be used for a SPDE. In a different spirit, Malliavin calculus is used in [31] to analyse adaptive schemes for the weak approximation of stochastic differential equations.

Our method is much closer to the method developped in [21]. There, the Malliavin calculus is also used to get rid of a stochastic integral which appears when writting down the weak error. However, it is done in a global way and the error is not decomposed as in the present article. A fundamental feature of Kohatsu-Higa's method is that the Kolmogorov equation is not used so that more general stochastic equation can be can considered. The solution does not need to be markovian. However, no SPDE have been considered with this method.

2. Preliminaries and main result

We consider the following stochastic partial differential equation written in an abstract form in a Hilbert space H with norm $|\cdot|$ and inner product $(\cdot,\cdot)$:

$$\left\{\begin{array}{l} dX = (AX + f(X))dt + \sigma(X)dW, \\ X(0) = x, \end{array}\right. \tag{2.1}$$

where the unknown X is a random process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ depending on $t > 0$ and on the initial data $x \in H$. The operator A is a negative self-adjoint operator on H with domain $D(A)$ and has a compact inverse. We assume that

$$\mathrm{Tr}((-A)^{-\alpha}) < \infty, \text{ for all } \alpha > 1/2. \tag{2.2}$$

We define classically the domain $D((-A)^{\beta})$, $\beta \in R$, of fractional powers of A and set

$$|x|_{\beta} = |(-A)^{\beta} x|, \; x \in D((-A)^{\beta}).$$

The nonlinear function f takes values in H and is assumed to be C^3 with bounded derivatives up to order 3. We denote by L_f a constant such that for $x, y \in H$

$$\begin{aligned} &|f(x)| \leq L_f(|x|+1), \\ &|f(x)-f(y)| \leq L_f|x-y|, \\ &|f'(x)-f'(y)|_{\mathcal{L}(H)} \leq L_f|x-y|. \end{aligned} \tag{2.3}$$

The noise is written in terms of a cylindrical Wiener process W on H (see [5]) associated to a filtration $(\mathcal{F}_t)_{t\geq 0}$. The nonlinear mapping acting on the noise maps H onto $\mathcal{L}(H)$, it is also assumed to be C^3 with bounded derivatives up to order 3. We denote by L_σ a constant satisfying

$$\begin{aligned} &|\sigma(x)|_{\mathcal{L}(H)} \leq L_\sigma(|x|+1), \\ &|\sigma(x)-\sigma(y)|_{\mathcal{L}(H)} \leq L_\sigma|x-y|. \end{aligned} \tag{2.4}$$

We need a stronger assumption on this mapping, we require

$$|\sigma''(x)\cdot(h,h)|_{\mathcal{L}(H))} \leq L_\sigma |h|^2_{-1/4},\ x\in H,\ h \in H. \tag{2.5}$$

Note that this implies a strong restriction on σ. (See Remark 2.3 below for some comments on this assumptions).

Recall that the cylindrical Wiener process can be written as

$$W = \sum_{\ell\in\mathbb{N}} \beta_\ell e_\ell$$

where here and in the following $(e_\ell)_{\ell\in\mathbb{N}}$ is any orthonormal basis of H and $(\beta_\ell)_{\ell\in\mathbb{N}}$ is an associated sequence of independent brownian motions. This series does not converge in H but in any larger Hilbert space U such that the embedding $H \subset U$ is Hilbert-Schmidt. Similarly, given a linear operator Φ from H to a possibly different Hilbert space K, the Wiener process $\Phi W = \sum_{\ell\in\mathbb{N}} \beta_\ell \Phi e_\ell$ is well defined in K provided $\Phi \in \mathcal{L}_2(H,K)$, the space of Hilbert-Schmidt operators from H to K. (See the definition just below).

Recall also that the stochastic integral $\int_0^T \Psi(s)dW(s)$ is defined as an element of K provided that Ψ is an adapted process with values in $\mathcal{L}_2(H,K)$ such that $\int_0^T |\psi(s)|^2_{\mathcal{L}_2(H,K)}ds < \infty$ *a.s.* (see [5]).

If $L \in \mathcal{L}(H)$ is a nuclear operator, $\mathrm{Tr}(L)$ denotes the trace of the operator L, i.e.

$$\mathrm{Tr}(L) = \sum_{i\geq 1}(Le_i, e_i) < +\infty.$$

It is well known that the previous definition does not depend on the choice of the Hilbertian basis. Moreover, the following properties hold for L nuclear and M bounded

$$\mathrm{Tr}(LM) = \mathrm{Tr}(ML), \tag{2.6}$$

and, if L is also positive,

$$\mathrm{Tr}(LM) \leq \mathrm{Tr}(L)\|M\|_{\mathcal{L}(H)}. \tag{2.7}$$

Hilbert-Schmidt operators play also an important role. An operator $L \in \mathcal{L}(H)$ is Hilbert-Schmidt if L^*L is a nuclear operator on H. We denote by $\mathcal{L}_2(H)$ the space of such operators. It is a Hilbert space for the norm

$$\|L\|_{\mathcal{L}_2(H)} = \left(\mathrm{Tr}(L^*L)\right)^{1/2} = \left(\mathrm{Tr}(LL^*)\right)^{1/2}.$$

It is classical that if $L \in \mathcal{L}_2(H)$, $M \in \mathcal{L}(H)$, $N \in \mathcal{L}(H)$ then $NLM \in \mathcal{L}_2(H)$ and

$$\|NLM\|_{\mathcal{L}_2(H)} \leq \|N\|_{\mathcal{L}(H)}\|L\|_{\mathcal{L}_2(H)}\|M\|_{\mathcal{L}(H)}. \tag{2.8}$$

See [5], appendix C, or [9] for more details on nuclear and Hilbert-Schmidt operators. Note that (2.2) implies that $(-A)^{-\beta}$ is Hilbert-Schmidt for any $\beta > 1/4$.

Our assumptions imply that for any $x \in H$, there exists a unique solution $X(t)$ to equation to (2.1) (see for instance [5], chapter 7). In the sequel, we often recall the dependence of the solution on the initial data by using the notation $X(t,x)$.

We approximate equation (2.1) by an implicit Euler schemes. Let $\Delta t = \frac{T}{N} > 0$ be a time step, we define the sequence $(X_k)_{k=0,\dots,N}$ by

$$\begin{cases} X_{k+1} = S_{\Delta t}X_k + \Delta t S_{\Delta t}f(X_k) + \sqrt{\Delta t}S_{\Delta t}\sigma(X_k)\chi_{k+1}, \\ X_0 = x. \end{cases} \tag{2.9}$$

We have set $\chi_{k+1} = (W((k+1)\Delta t) - W(k\Delta t))/\sqrt{\Delta t}$. The operators $S_{\Delta t}$ is defined by

$$S_{\Delta t} = (I - \Delta t A)^{-1}.$$

This is the classical fully implicit Euler scheme. It will be convenient to use the integral form of (2.1)

$$X(t) = S(t)x + \int_0^t S(t-s)f(X(s))ds + \int_0^t S(t-s)\sigma(X(s))dW(s),\ t \geq 0, \tag{2.10}$$

where $S(t) = e^{tA}$ is the semigroup generated by A. Similarly, (2.9) can be rewritten as

$$X_k = S_{\Delta t}^k x + \Delta t \sum_{\ell=0}^{k-1} S_{\Delta t}^{k-\ell} f(X_\ell) + \sqrt{\Delta t}\sum_{\ell=0}^{k-1} S_{\Delta t}^{k-\ell}\sigma(X_\ell)\chi_{\ell+1}. \tag{2.11}$$

It will be convenient in the following to use the notation:

$$t_k = k\Delta t,\ k = 0,\dots,N.$$

The following inequalities are classical and easily proved using the spectral decomposition of A

$$\left|(-A)^\beta S_{\Delta t}^k\right|_{\mathcal{L}(H)} \leq c t_k^{-\beta}, \quad k \geq 1, \quad \beta \in [0,1]. \tag{2.12}$$

$$\left|(-A)^\beta S(t)\right|_{\mathcal{L}(H)} \leq c t^{-\beta}, \quad t > 0, \quad \beta \geq 0. \tag{2.13}$$

$$\left|(-A)^\beta S_{\Delta t}\right|_{\mathcal{L}(H)} \leq c \Delta t^{-\beta}, \quad \beta \in [0,1]. \tag{2.14}$$

$$\left|(-A)^{-\beta}\left(I - S_{\Delta t}\right)\right|_{\mathcal{L}(H)} \leq c\Delta t^{\beta}, \quad \beta \in [0,1]. \tag{2.15}$$

Note that in (2.11) and (2.9), the noise term makes sense in H. Indeed, by (2.14), (2.2) and (2.8), we know that $S_{\Delta t}$ is a Hilbert-Schmidt operator on H.

We are interested in the approximation of the law of the solution of (2.1). More precisely, we wish to prove an estimate on the error committed when approximating $\mathbb{E}(\varphi(X(T,x)))$ by $\mathbb{E}(\varphi(X_N(x)))$. The function φ is a smooth function on H.

In all the article, we use the notation $D\varphi(x)$ for the differential of a C^1 function on H at the point x. If $\varphi:\ H\mapsto K$, where K is another Hilbert space, $D\varphi(x)\in\mathcal{L}(H,K)$ the space of continuous linear operator from H to K. When $K=\mathbb{R}$, we identify the differential with the gradient thanks to Riesz identification theorem. We use the same notation and have the identity for $x,h\in H$:

$$D\varphi(x).h=(D\varphi(x),h).$$

Similarly, if $\varphi\in C^2(H,\mathbb{R})$, $D^2\varphi(x)$ is a bilinear operator from $H\times H$ to $\mathbb{R}$ and can be identified with a linear operator on H through the identity:

$$D^2\varphi(x).(h,k)=(D^2\varphi(x)h,k),\ x,h,k\in H.$$

Sometimes, we also use the notations φ', φ'' instead of $D\varphi$ or $D^2\varphi$.

Given two Banach spaces K_1 and K_2, we denote by $\|\cdot\|_k$ the norm on $C_b^k(K_1,K_2)$, the space of k times continuously differentiable mapping from K_1 to K_2 with derivatives bounded up to order k.

We use Malliavin calculus in the course of the proof. We now recall the basic definitions. (See [28]). Given a smooth real valued function F on H^n and $\psi_1,\dots,\psi_n\in L^2(0,T,H)$, the Malliavin derivative of the smooth random variable $F(\int_0^T(\psi_1(s),dW(s)),\dots,\int_0^T(\psi_n(s),dW(s)))$ at time s in the direction $h\in H$ is given by

$$\begin{aligned}&D_s^h\left[F\left(\int_0^T(\psi_1(s),dW(s)),\dots,\int_0^T(\psi_n(s),dW(s))\right)\right]\\&=\sum_{i=1}^n\partial_iF\left(\int_0^T(\psi_1(s),dW(s)),\dots,\int_0^T(\psi_n(s),dW(s))\right)(\psi_i(s),h)\,.\end{aligned}$$

We also define the process DF by $(DF(s),h)=D_s^hF$. It can be shown that D defines a closable operator with values in $L^2(\Omega,L^2(0,T,H))$ and we denote by $\mathbb{D}^{1,2}$ the closure of the set of smooth random variables as above for the topology defined by the norm

$$\|F\|_{\mathbb{D}^{1,2}}=\left(\mathbb{E}(|F|^2)+\mathbb{E}(\int_0^T|D_sF|^2ds\right)^{1/2}.$$

We define similarly the Malliavin derivative of random variables taking values in H. If $G=\sum_{i\in\mathbb{N}}F_ie_i\in L^2(\Omega,H)$ where $F_i\in\mathbb{D}^{1,2}$ for all $i\in\mathbb{N}$ and $\sum_{i\in\mathbb{N}}\int_0^T|D_sF_i|^2ds<\infty$, we set $D_s^hG=\sum_{i\in\mathbb{N}}D_s^hF_ie_i$, $D_sG=\sum_{i\in\mathbb{N}}D_sF_ie_i$. We define $\mathbb{D}^{1,2}(H)$ as the set of such random variables.

When $h=e_m$, we write $D^{e_m}=D^m$.

The chain rule is valid and given $u\in C_b^1(\mathbb{R})$, $F\in\mathbb{D}^{1,2}$ then $u(F)\in\mathbb{D}^{1,2}$ and $D(u(F))=u'(F)DF$. Also if $G=\sum_{i\in\mathbb{N}}F_ie_i\in\mathbb{D}^{1,2}(H)$ and $u\in C_b^1(H,\mathbb{R})$ then $u(G)\in\mathbb{D}^{1,2}$ and $D(u(G))=Du(G).DG=(Du,DG)$, or equivalently $D_s^h(u(G))=\sum_{i\in\mathbb{N}}\partial_iuD_s^hF_i=(Du,D_s^hG)$.

Note that as already mentionned, we identify the differential of a function in $C^1(H,\mathbb{R})$ with its gradient.

For $F \in \mathbb{D}^{1,2}$ and $\psi \in L^2(\Omega \times [0,T]; H)$ such that $\psi(t) \in \mathbb{D}^{1,2}$ for all $t \in [0,T]$ and $\int_0^T \int_0^T |D_s\psi(t)|^2 dsdt < \infty$, we have the integration by part formula:

$$\mathbb{E}\left(F \int_0^T (\psi(s), dW(s)) \right) = \mathbb{E}\left(\int_0^T (D_s F, \psi(s)) ds \right),$$

where the stochastic integral is a Skohorod integral which is in fact defined by duality. In this article, we only need to consider the Skohorod integral of adapted processes in which case it corresponds with the Itô integral. Moreover, the integration by part formula above holds for $F \in \mathbb{D}^{1,2}$ and $\psi \in L^2(\Omega \times [0,T]; H)$ when ψ is an adapted process. Recall that if F is $\mathcal{F}_t$ measurable then $D_s F = 0$ for $s \geq t$.

We will often use the following form of the integration by part formula whose proof is left to the reader.

Lemma 2.1. *Let $F \in \mathbb{D}^{1,2}(H)$, $u \in C_b^2(H)$ and $\psi \in L^2(\Omega \times [0,T], \mathcal{L}_2(H))$ be an adapted process then*

$$\begin{aligned} \mathbb{E}\left(Du(F) \cdot \int_0^T \psi(s) dW(s) \right) &= \mathbb{E}\left(\sum_{m \in \mathbb{N}} \int_0^T D^2u(F) \cdot (D_s^m F, \psi(s) e_m) ds \right) \\ &= \mathbb{E}\left(\int_0^T \mathrm{Tr}\left(\psi^*(s) D^2 u(F) D_s F \right) ds \right). \end{aligned}$$

Also we remark that this Lemma remains valid if u is not assumed to be bounded but only $u \in C^2(H)$ provided the expectations and the integral above are well defined. This is easily seen by approximation of u by bounded functions.

We now state our main result.

Theorem 2.2. *Assume that f and σ are C_b^3 functions from H to H and $\mathcal{L}(H)$ and that σ satisfies* (2.4)*, then for any $x \in H$, $T > 0$, $\varepsilon > 0$, the Euler Scheme* (2.9) *satisfies the following weak error estimate*

$$|\mathbb{E}(\varphi(X(T,x))) - \mathbb{E}(\varphi(X_N))| \leq C(T, |\varphi|_{C_b^3}, |x|, \varepsilon) \Delta t^{1/2-\varepsilon}, \quad \varphi \in C_b^3(H).$$

Remark 2.3. *Assumption* (2.4) *is quite restrictive. It is void for an additive noise or a noise of the form $BX\,dW$ where B is a linear operator from H to $\mathcal{L}(H)$. Otherwise, it implies that the noise is a perturbation of such noise. An example of a noise satisfying this is*

$$\sigma(x) = Bx + \tilde{\sigma}((-A)^{-1/4} x)$$

where $B \in \mathcal{L}(H)$ and $\tilde{\sigma} : H \to \mathcal{L}(H)$ is a C^3 function with derivatives bounded up to order 3. This assumption is crucial in our proof. It is used in essential way in Lemma 4.5 which is used at many points of the proof.

Apart from this point, our result is optimal. If the noise is assumed to satisfied some non degeneracy assumptions, the smothness assumption on the test function φ can be weakened. This will be investigated in a future work.

In all the article, C or c denote constants which may depend on A, f, σ, Q or T but not on Δt. Their value may change from one line to another. The initial data x is fixed and the constant may also depend on $|x|$. Note also that we assume that $\Delta t \leq 1$, we could also assume

$\Delta t \le \Delta t_0$ for some $\Delta t_0 > 0$. In this case, the different constants would depend on Δt_0. Finally, ε is a small positive number.

3. Proof of the main result

The proof uses different tools from stochastic calculus such as Itô formula, Kolmogorov equations, Malliavin calculus. Sometimes, it may be very lengthy and technical to justify rigorously their use in infinite dimension. We avoid these tedious justifications by using Gakerkin approximations. We replace equation (2.1) by the finite dimensional stochastic equation

$$dX_m = (AX_m + f_m(X_m))dt + \sigma_m(X_m)dW, \quad X_m(0) = P_m$$

where P_m is the eigenprojector on the m first eignevectors of A, $f_m(x) = P_m f(x)$, $\sigma_m(x) = P_m\sigma(x)P_m$. It is not difficult to prove that X_m converges to X in various senses.

Similarly, we replace the discrete unknown X_k by a finite dimensional sequence defined in an obvious way.

We prove the result for these finite dimensional objects with constants that do not depend on the dimension m. It is then easy to deduce the result for our infinite dimensional equation.

In order to lighten the notation, we omit to explicit the dependence on m below and write X, f, σ instead of X_m, f_m, σ_m.

Step 1: We first define a continuous interpolation of the discrete unknown.

We rewrite (2.9) as follows:

$$X_{k+1} = X_k + \int_{t_k}^{t_{k+1}} A_{\Delta t}X_k + S_{\Delta t}f(X_k)ds + \int_{t_k}^{t_{k+1}} S_{\Delta t}\sigma(X_k)dW(s)$$

where $A_{\Delta t} = S_{\Delta t}A$. Note that $A_{\Delta t}$ is in fact a Yosida regularization of A and is a bounded operator:

$$|A_{\Delta t}|_{\mathcal{L}(H)} \le c\Delta t^{-1}. \tag{3.1}$$

It is then natural to define $\tilde{X}$ on $[0,T]$ by

$$\tilde{X}(t) = X_k + \int_{t_k}^{t} A_{\Delta t}X_k + S_{\Delta t}f(X_k)ds + \int_{t_k}^{t} S_{\Delta t}\sigma(X_k)dW(s), \quad t \in [t_k, t_{k+1}). \tag{3.2}$$

Clearly, $\tilde{X}$ is a continuous and adapted process. Given a smooth function G on $[0,T]\times H$, Itô formula implies for $t \in [t_k, t_{k+1})$ (see [5]):

$$\begin{aligned} G(t,\tilde{X}(t)) &= G(t_k,\tilde{X}(t_k)) + \int_{t_k}^{t} \frac{dG}{dt}(s,\tilde{X}(s)) + L_{k,\Delta t}G(s,\tilde{X}(s))ds \\ &\quad + \int_{t_k}^{t} (DG(s,\tilde{X}(s)), \sigma(X_k)dW(s)). \end{aligned} \tag{3.3}$$

Where for $\psi \in C^2(H,\mathbb{R})$

$$L_{k,\Delta t}\psi(x) = \frac{1}{2}\mathrm{Tr}\left\{(S_{\Delta t}\sigma(X_k))\,(S_{\Delta t}\sigma(X_k))^*\,D^2\psi(x)\right\} + (A_{\Delta t}X_k + S_{\Delta t}f(X_k), D\psi(x)).$$

Step 2: Decomposition of the error.

Let us define

$$u(t,x)=\mathbb{E}(\varphi(X(t,x))),\ t\in[0,T]. \tag{3.4}$$

Then the weak error at time T is equal to

$$\begin{aligned} u(T,x)-\mathbb{E}(\varphi(X_N)) &= \mathbb{E}(u(T,x)-u(0,X_N) \\ &= \sum_{k=0}^{N-1}\mathbb{E}\left(u(T-t_k,X_k)-u(T-t_{k+1},X_{k+1})\right). \end{aligned} \tag{3.5}$$

It is well known that u is a solution to the forward Kolmogorov equation:

$$\begin{aligned} \frac{du}{dt}(t,x) &= Lu(t,x) \\ &= \frac{1}{2}\mathrm{Tr}\{\sigma(x)\sigma^*(x)D^2u(t,x)\}+(Ax+f(x),Du(t,x)). \end{aligned} \tag{3.6}$$

Therefore, Itô formula (3.3) implies

$$\mathbb{E}(u(T-t_{k+1},X_{k+1}))=\mathbb{E}(u(T-t_k,X_k))+\mathbb{E}\int_{t_k}^{t_{k+1}}L_{k,\Delta t}u(T-t,\tilde X(t))-Lu(T-t,\tilde X(t))dt.$$

The first term in (3.5) will be treated separately and we decompose the error as follows

$$u(T,x)-\mathbb{E}(\varphi(X_N))=u(T,x)-\mathbb{E}(u(T-\Delta t,X_1))+\sum_{k=1}^{N-1}a_k+b_k+c_k. \tag{3.7}$$

Where

$$a_k=\mathbb{E}\int_{t_k}^{t_{k+1}}\Big(A\tilde X(t)-A_{\Delta t}X_k,Du(T-t,\tilde X(t))\Big)\,dt,$$

$$b_k=\mathbb{E}\int_{t_k}^{t_{k+1}}\Big(f(\tilde X(t))-S_{\Delta t}f(X_k),Du(T-t,\tilde X(t))\Big)\,dt,$$

$$c_k=\frac{1}{2}\mathbb{E}\int_{t_k}^{t_{k+1}}\mathrm{Tr}\left\{\left[\sigma(\tilde X(t))\sigma^*(\tilde X(t))-\left(S_{\Delta t}\sigma(X_k)\right)\left(S_{\Delta t}\sigma(X_k)\right)^*\right]D^2u(T-t,\tilde X(t))\right\}dt.$$

In the next steps, we estimate separately the different terms in (3.7).

Step 3: Estimate of $u(T,x)-\mathbb{E}(u(T-\Delta t,X_1))$.

By the Markov property

$$u(T,x)=\mathbb{E}(\varphi(X(T,x)))=\mathbb{E}(u(T-\Delta t,X(\Delta t))).$$

Therefore, by Lemma 4.4, for any $\varepsilon>0$,

$$|u(T,x)-\mathbb{E}(u(T-\Delta t,X_1))|\le c(T-\Delta t)^{-1/2+\varepsilon}\|\varphi\|_1\mathbb{E}\left(|X(\Delta t)-X_1|_{-1/2+\varepsilon}\right).$$

Moreover

$$\begin{aligned} X(\Delta t)-X_1 &= (S(\Delta t)-S_{\Delta t})x+\textstyle\int_0^{\Delta t}S(t-s)f(X(s,x))ds-\Delta t S_{\Delta t}f(x) \\ &\quad+\textstyle\int_0^{\Delta t}S(t-s)\sigma(X(s,x))dW(s)-\sqrt{\Delta t}S_{\Delta t}\sigma(x)\chi_1. \end{aligned}$$

It is easy to prove that

$$|(-A)^{-1/2+\varepsilon}\left(S(\Delta t)-S_{\Delta t}\right)|_{\mathcal{L}(H)}\leq c\Delta t^{1/2-\varepsilon}.$$

Since $(S(t))_{t\geq 0}$ is a contraction semigroup and $|(-A)^{-1/2+\varepsilon}\cdot|\leq c|\cdot|$, we have by (2.3) and Lemma 4.2

$$\mathbb{E}\left|\int_0^{\Delta t}S(t-s)f(X(s,x))ds\right|_{-1/2+\varepsilon}\leq\Delta tL_f\mathbb{E}(\sup_{s\in[0,\Delta t]}|X(s,x)|+1)\leq c\Delta t(|x|+1).$$

Similarly

$$|\Delta tS_{\Delta t}f(x)|_{-1/2+\varepsilon}\leq c\Delta t(|x|+1).$$

We then have

$$\begin{aligned}&\mathbb{E}\left(|\int_0^{\Delta t}S(t-s)\sigma(X(s,x))dW(s)|^2_{-1/2+\varepsilon}\right)\\&=\mathbb{E}\left(\int_0^{\Delta t}|(-A)^{-1/2+\varepsilon}S(t-s)\sigma(X(s,x))|^2_{\mathcal{L}_2(H)}ds\right)\\&\leq\mathbb{E}\left(\int_0^{\Delta t}|(-A)^{-1/2+\varepsilon}|^2_{\mathcal{L}_2(H)}|S(t-s)|^2_{\mathcal{L}(H)}|\sigma(X(s,x))|^2_{\mathcal{L}(H)}ds\right)\end{aligned}$$

and by (2.2), (2.4), Lemma 4.2

$$\mathbb{E}\left(|\int_0^{\Delta t}S(t-s)\sigma(X(s,x))dW(s)|^2_{-1/2+\varepsilon}\right)\leq c\Delta t(|x|+1).$$

Similarly

$$\mathbb{E}\left(|\sqrt{\Delta t}S_{\Delta t}\sigma(x)\chi_1|^2\right)\leq c\Delta t(|x|+1).$$

Gathering these estimate and using Cauchy-Schwartz inequality , we obtain

$$|u(T,x)-\mathbb{E}(u(T-\Delta t,X_1))|\leq c(T-\Delta t)^{-1/2+\varepsilon}\Delta t^{-1/2+\varepsilon}\leq c\Delta t^{1/2-\varepsilon}\tag{3.8}$$

where, as mentionned above, the constant is allowed to depend on T, x, φ, f, σ ...

Step 4: Estimate of a_k, $k\geq 1$.

We split a_k as follows:

$$a_k=a_k^1+a_k^2$$

with

$$a_k^1=\mathbb{E}\int_{t_k}^{t_{k+1}}\left((A-A_{\Delta t})X_k,Du(T-t,\tilde{X}(t))\right)dt,$$

$$a_k^2=\mathbb{E}\int_{t_k}^{t_{k+1}}\left(A(\tilde{X}(t)-X_k),Du(T-t,\tilde{X}(t))\right)dt.$$

Note that $A_{\Delta t}-A=\theta\Delta tS_{\Delta t}A^2$. By Lemma 4.4 below, we know that $Du(T-t,\tilde{X}(t))$ is in $D((-A)^\gamma)$ for $\gamma<1/2$ and it is easy to see that X_k belongs to $D((-A)^\delta)$ for $\delta<1/4$. It is impossible to compensate the presence of A^2 by such arguments. The idea is to recall (2.11) and to observe that the irregularity of X_k is contained in the stochastic integral. We thus further decompose a_k^1 in three terms according to (2.11). The first two terms are easy to treat. The third one involves the stochastic integral and is estimated thanks to Malliavin calculus. We set

$$a_k^{1,1} = -\theta\Delta t\mathbb{E}\int_{t_k}^{t_{k+1}} \left(S_{\Delta t}A^2S_{\Delta t}^k x, Du(T-t,\tilde{X}(t))\right)dt,$$

$$a_k^{1,2} = -\theta\Delta t\mathbb{E}\int_{t_k}^{t_{k+1}} \left(S_{\Delta t}A^2\Delta t\sum_{\ell=0}^{k-1} S_{\Delta t}^{k-\ell}f(X_\ell), Du(T-t,\tilde{X}(t))\right)dt,$$

$$a_k^{1,3} = -\theta\Delta t\mathbb{E}\int_{t_k}^{t_{k+1}} \left(S_{\Delta t}A^2\sqrt{\Delta t}\sum_{\ell=0}^{k-1} S_{\Delta t}^{k-\ell}\sigma(X_\ell)\chi_{\ell+1}, Du(T-t,\tilde{X}(t))\right)dt,$$

so that

$$a_k^1 = a_k^{1,1} + a_k^{1,2} + a_k^{1,3}.$$

By (2.14), (2.12) and Lemma 4.4, we have for $k = 1,\dots,N-2$ and $\varepsilon > 0$

(3.9)

$$\begin{aligned}|a_k^{1,1}| &\le c\Delta t\mathbb{E}\int_{t_k}^{t_{k+1}} |S_{\Delta t}(-A)^{1/2+2\varepsilon}|_{\mathcal{L}(H)}|(-A)^{1-\varepsilon}S_{\Delta t}^k|_{\mathcal{L}(H)}|(-A)^{1/2-\varepsilon}Du(T-t,\tilde{X}(t))|\,|x|dt \\ &\le c\Delta t^{1/2-2\varepsilon}t_k^{-1+\varepsilon}\int_{t_k}^{t_{k+1}}(T-t)^{-(1/2-\varepsilon)}dt.\end{aligned}$$

The estimate of $a_k^{1,2}$ is similar. We have by (2.3), (2.12)

$$\begin{aligned}\left|\Delta t(-A)^{1-\varepsilon}\sum_{\ell=0}^{k-1} S_{\Delta t}^{k-\ell}f(X_\ell)\right| &\le L_f\Delta t\sum_{\ell=0}^{k-1}\left|(-A)^{1-\varepsilon}S_{\Delta t}^{k-\ell}\right|_{\mathcal{L}(H)}(|X_\ell|+1) \\ &\le c\Delta t\sum_{\ell=0}^{k-1} t_{k-\ell}^{-1+\varepsilon}(|X_\ell|+1).\end{aligned}$$

Since

$$\Delta t\sum_{\ell=0}^{k-1} t_{k-\ell}^{-1+\varepsilon} \le \varepsilon^{-1}T^\varepsilon,$$

we deduce thanks to Lemma 4.4 and Lemma 4.1

$$\begin{aligned}|a_k^{1,2}| &\le c\Delta t\int_{t_k}^{t_{k+1}}\left|S_{\Delta t}(-A)^{1/2+2\varepsilon}\right|_{\mathcal{L}(H)}(T-t)^{-1/2+\varepsilon}dt \\ &\le c\Delta t^{1/2-2\varepsilon}\textstyle\int_{t_k}^{t_{k+1}}(T-t)^{-(1/2-\varepsilon)}dt.\end{aligned} \tag{3.10}$$

To treat $a_k^{1,3}$, we first rewrite it in terms of a stochastic integral and then use Lemma 2.1

$$\begin{aligned} a_k^{1,3} &= \theta\Delta t\mathbb{E}\int_{t_k}^{t_{k+1}}\left(\int_0^{t_k} S_{\Delta t}A^2 S_{\Delta t}^{k-\ell_s}\sigma(X_{\ell_s})dW(s), Du(T-t,\tilde X(t))\right)dt \\ &= \theta\Delta t\mathbb{E}\int_{t_k}^{t_{k+1}}\int_0^{t_k}\mathrm{Tr}\left\{\sigma^*(X_{\ell_s})S_{\Delta t}A^2S_{\Delta t}^{k-\ell_s}D^2u(T-t,\tilde X(t))D_s\tilde X(t)\right\}ds\ dt \end{aligned}$$

where $\ell_s = [s/\Delta t]$ is the integer part of $s/\Delta t$. By the chain rule and (3.2), we have for $s\in[0,t_k]$, $h\in H$, $t\in[t_k,t_{k+1})$,

$$D_s^h\tilde X(t) = D_s^hX_k + \int_{t_k}^t A_{\Delta t}D_s^hX_k + S_{\Delta t}f'(X_k)\cdot D_s^hX_kds + \int_{t_k}^t S_{\Delta t}\left(\sigma'(X_k)\cdot D_s^hX_k\right)dW(s).$$

Fro $\beta<1/4$, we have, by (2.14), (2.2), (2.4), (2.8)

$$\begin{aligned} &\mathbb{E}\left(\left|\int_{t_k}^t S_{\Delta t}\left(\sigma'(X_k)\cdot D_s^hX_k\right)dW(s)\right|_\beta^2\right) \\ &= \mathbb{E}\left(\int_{t_k}^t\left|(-A)^\beta S_{\Delta t}\left(\sigma'(X_k)\cdot D_s^hX_k\right)\right|_{\mathcal{L}_2(H)}^2ds\right) \\ &\le \mathbb{E}\left(\int_{t_k}^t\left|(-A)^{-1/4-\varepsilon}\right|_{\mathcal{L}_2(H)}^2\left|(-A)^{\beta+1/4+\varepsilon}S_{\Delta t}\right|_{\mathcal{L}(H)}^2\left|\sigma'(X_k)\cdot D_s^hX_k\right|_{\mathcal{L}(H)}^2ds\right) \\ &\le c\Delta t^{1/2-2\beta-2\varepsilon}\mathbb{E}\left(|D_s^hX_k|^2\right). \end{aligned}$$

We then use (3.1), (2.3) to bound the other terms above and obtain thanks to Poincaré inequality

$$\mathbb{E}\left(|D_s^h\tilde X(t)|_\beta^2\right)\le c\mathbb{E}\left(|D_s^hX_k|_\beta^2\right),\quad s\in[0,t_k],\quad t\in[t_k,t_{k+1}). \tag{3.11}$$

By Lemma 4.3, we obtain for $\beta<1/4$

$$\mathbb{E}\left(\left|(-A)^\beta D_s\tilde X(t)\right|_{\mathcal{L}(H)}^2\right)\le ct_{k-\ell_s}^{-2\beta}.$$

We are now ready to conclude the estimate of $a_k^{1,3}$. We choose $\varepsilon>0$ and write thanks to (2.4), (2.14), (2.12), Lemma 4.5 and (2.2)

$$\begin{aligned} |a_k^{1,3}| \le\ & \theta\Delta t\mathbb{E}\int_{t_k}^{t_{k+1}}\int_0^{t_k}|\sigma^*(X_{\ell_s})|_{\mathcal{L}(H)}\left|S_{\Delta t}A^{1/2+2\varepsilon}\right|_{\mathcal{L}(H)}\left|(-A)^{1-3\varepsilon/2}S_{\Delta t}^{k-\ell_s}\right|_{\mathcal{L}(H)} \\ &\times\left|(-A)^{1/2-\varepsilon/2}D^2u(T-t,\tilde X(t))(-A)^{1/2-\varepsilon/2}\right|_{\mathcal{L}(H)}\mathrm{Tr}\left\{(-A)^{-1/2-\varepsilon/2}\right\}\left|(-A)^\varepsilon D_s\tilde X(t)\right|_{\mathcal{L}(H)}ds\ dt \\ &\le c\Delta t\mathbb{E}\int_{t_k}^{t_{k+1}}\int_0^{t_k}\Delta t^{-1/2-2\varepsilon}t_{k-\ell_s}^{-1+3\varepsilon/2}(T-t)^{-1+\varepsilon}t_{k-\ell_s}^{-\varepsilon}ds\ dt. \end{aligned}$$

Since $\int_0^{t_k} t_{k-\ell_s}^{-1+\varepsilon/2} ds \leq \frac{2}{\varepsilon} T^{\varepsilon/2}$, we deduce

$$|a_k^{1,3}| \leq c\Delta t^{1/2-2\varepsilon} \int_{t_k}^{t_{k+1}} (T-t)^{-1+\varepsilon} dt. \tag{3.12}$$

Gathering (3.9), (3.10) and (3.12), we obtain for $k = 1, \dots, N-1$

$$|a_k^1| \leq c\Delta t^{1/2-2\varepsilon} \left(t_k^{-1+\varepsilon} + 1\right) \left(\int_{t_k}^{t_{k+1}} (T-t)^{-1+\varepsilon} dt + 1 \right). \tag{3.13}$$

We now estimate a_k^2. Let us set

$$a_k^{2,1} = \mathbb{E} \int_{t_k}^{t_{k+1}} (t - t_k) \Big(AA_{\Delta t} X_k, Du(T-t, \tilde{X}(t)) \Big) \, dt,$$

$$a_k^{2,2} = \mathbb{E} \int_{t_k}^{t_{k+1}} (t - t_k) \Big(AS_{\Delta t} f(X_k), Du(T-t, \tilde{X}(t)) \Big) \, dt,$$

$$a_k^{2,3} = \mathbb{E} \int_{t_k}^{t_{k+1}} \int_{t_k}^{t} \Big(AS_{\Delta t} \sigma(X_k) dW(s), Du(T-t, \tilde{X}(t)) \Big) \, dt,$$

so that thanks to (3.2), we have $a_k^2 = a_k^{2,1} + a_k^{2,2} + a_k^{2,3}$. The first term $a_k^{2,1}$ is similar to a_k^1 above and is majorized in the same way

$$|a_k^{2,1}| \leq c\Delta t^{1/2-2\varepsilon} \left(t_k^{-1+\varepsilon} + 1\right) \left(\int_{t_k}^{t_{k+1}} (T-t)^{-1+\varepsilon} dt + 1 \right). \tag{3.14}$$

for $k = 1, \dots, N-1$. The second one is not difficult to treat, we have using similar arguments as above

$$\begin{aligned} |a_k^{2,2}| &\leq c\Delta t |(-A)^{1/2+\varepsilon} S_{\Delta t}|_{\mathcal{L}(H)} \mathbb{E}(|f(X_k)|) \int_{t_k}^{t_{k+1}} (T-t)^{-(1/2-\varepsilon)} dt \\ &\leq c\Delta t^{1/2-\varepsilon} \int_{t_k}^{t_{k+1}} (T-t)^{-(1/2-\varepsilon)} dt \end{aligned} \tag{3.15}$$

for $k = 1, \dots, N-1$. The estimate of $a_k^{2,3}$ requires the use of Lemma 2.1. It implies

$$a_k^{2,3} = \mathbb{E} \int_{t_k}^{t_{k+1}} \int_{t_k}^{t} \operatorname{Tr} \Big\{ \sigma^*(X_k) S_{\Delta t} A D^2 u(T-t, \tilde{X}(t)) D_s \tilde{X}(t) \Big\} \, ds \, dt.$$

Since, X_k is $\mathcal{F}_{t_k}$ measurable, we have from (3.2)

$$D_s \tilde{X}(t) = S_{\Delta t} \sigma(X_k) \quad s \in (t_k, t_{k+1}], \quad t_k \leq s \leq t < t_{k+1}. \tag{3.16}$$

It follows, thanks to (2.4), (2.14), (2.2) and Lemma 4.5,

$$a_k^{2,3} = \mathbb{E}\int_{t_k}^{t_{k+1}} (t-t_k)\mathrm{Tr}\left\{\sigma^*(X_k)S_{\Delta t}AD^2u(T-t,\tilde{X}(t))S_{\Delta t}\sigma(X_k)\right\}dt$$

$$\leq c\Delta t\mathbb{E}\int_{t_k}^{t_{k+1}} |\sigma(X_k)|_{\mathcal{L}(H)}|S_{\Delta t}(-A)^{1/2+\varepsilon/2}|_{\mathcal{L}(H)}|(-A)^{1/2-\varepsilon/2}D^2u(T-t,\tilde{X}(t))(-A)^{1/2-\varepsilon/2}|_{\mathcal{L}(H)}$$

$$\times\mathrm{Tr}((-A)^{-1/2-\varepsilon/2})|(-A)^{\varepsilon}S_{\Delta t}|_{\mathcal{L}(H)}|\sigma(X_k)|_{\mathcal{L}(H)}dt$$

$$\leq c\Delta t^{1/2-3\varepsilon/2}\int_{t_k}^{t_{k+1}}(T-t)^{-1+\varepsilon}dt$$

for $k=1,\dots,N-1$. Finally, we obtain

$$|a_k^2| \leq c\Delta t^{1/2-2\varepsilon}(t_k^{-1+\varepsilon}+1)\left(\int_{t_k}^{t_{k+1}}(T-t)^{-1+\varepsilon}dt+1\right) \tag{3.17}$$

for $k=1,\dots,N-1$. Together with (3.13) this yields the estimate of a_k

$$|a_k| \leq c\Delta t^{1/2-2\varepsilon}(t_k^{-1+\varepsilon}+1)\left(\int_{t_k}^{t_{k+1}}(T-t)^{-1+\varepsilon}dt+1\right).$$

It follows easily

$$\sum_{k=1}^{N-1}|a_k| \leq c\Delta t^{1/2-2\varepsilon}. \tag{3.18}$$

Step 5: Estimate of b_k.

This term seems easier to treat since we do not have the unbounded operator A. However, since it involves the nonlinear term, we need to use Ito formula (3.3) to control $f(\tilde{X}(t))-f(X_k)$, this introduces many terms. For some of them we again use Malliavin integration by parts.

First, we get rid of $S_{\Delta t}$. We have thanks to (2.15), (2.3), Lemma 4.4 and Lemma 4.1:

$$\begin{aligned}
b_k^1 &= \mathbb{E}\int_{t_k}^{t_{k+1}}\Big((I-S_{\Delta t})\,f(X_k), Du(T-t,\tilde{X}(t))\Big)\,dt \\
&\leq c\mathbb{E}\int_{t_k}^{t_{k+1}}(1+|X_k|)|(-A)^{-1/2+\varepsilon}(I-S_{\Delta t})|_{\mathcal{L}(H)}|(-A)^{1/2-\varepsilon}Du(T-t,\tilde{X}(t))|dt \\
&\leq c\Delta t^{1/2-\varepsilon}\mathbb{E}\int_{t_k}^{t_{k+1}}(T-t)^{-1/2+\varepsilon}dt
\end{aligned}$$

for $k=0,\dots,N-1$. We now estimate

$$\begin{aligned}
b_k^2 &= b_k - b_k^1 \\
&= \mathbb{E}\int_{t_k}^{t_{k+1}}\Big(f(\tilde{X}(t)-f(X_k), Du(T-t,\tilde{X}(t))\Big)\,dt \\
&= \mathbb{E}\int_{t_k}^{t_{k+1}}\sum_{i\in\mathbb{N}}(f_i(\tilde{X}(t)-f_i(X_k))\partial_i u(T-t,\tilde{X}(t))dt,
\end{aligned}$$

where $f_i = (f, e_i)$ and $\partial_i = (D\cdot, e_i)$. We choose $(e_i)_{i\in\mathbb{N}}$ as the orthonormal basis of eigenvectors of A. By (3.3), we have for $i \in \mathbb{N}$

$$\begin{aligned} f_i(\tilde{X}(t) &= f_i(X_k) + \int_{t_k}^t \frac{1}{2}\mathrm{Tr}\left\{(S_{\Delta t}\sigma(X_k))(S_{\Delta t}\sigma(X_k))^* D^2 f_i(\tilde{X}(s))\right\} ds \\ &\quad + \int_{t_k}^t \Big(A_{\Delta t}X_k + S_{\Delta t}f(X_k), Df_i(\tilde{X}(s))\Big)\, ds + \int_{t_k}^t \Big(Df_i(\tilde{X}(s)), \sigma(X_k)\Big)\, dW(s). \end{aligned}$$

With obvious notations, this defines the decomposition

$$b_k^2 = b_k^{2,1} + b_k^{2,2} + b_k^{2,3} + b_k^{2,4}.$$

To treat the first term, we rewrite it as follows[1]:

$$\begin{aligned} b_k^{2,1} &= \frac{1}{2}\mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^t \sum_{i\in\mathbb{N}} \mathrm{Tr}\left\{(S_{\Delta t}\sigma(X_k))(S_{\Delta t}\sigma(X_k))^* D^2 f_i(\tilde{X}(s))\right\} \partial_i u(T-t, \tilde{X}(t)) ds\, dt \\ &= \frac{1}{2}\mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^t \mathrm{Tr}\left\{(S_{\Delta t}\sigma(X_k))(S_{\Delta t}\sigma(X_k))^* \mathcal{A}(s,t)\right\} ds\, dt \end{aligned}$$

where $\mathcal{A}(s,t) \in \mathcal{L}(H)$ is defined by

$$\begin{aligned} (\mathcal{A}(s,t)h, k) &= \sum_{i\in\mathbb{N}} D^2 f_i(\tilde{X}(s)).(h,k)\partial_i u(T-t, \tilde{X}(t)) \\ &= \Big(D^2 f(\tilde{X}(s)).(h,k), Du(T-t, \tilde{X}(t))\Big),\ h, k \in H. \end{aligned}$$

Obviously

$$|\mathcal{A}(s,t)|_{\mathcal{L}(H)} \le \left|D^2 f(\tilde{X}(s))\right|_{\mathcal{L}^2(H\times H, H)} \left|Du(T-t, \tilde{X}(t))\right|,$$

where $\mathcal{L}^2(H \times H, H)$ denotes the space of bilinear operators from $H \times H$ to H. By (2.3) and Lemma 4.4, we deduce:

$$|\mathcal{A}(s,t)|_{\mathcal{L}(H)} \le c.$$

Then, we write thanks to (2.4), (2.14), (2.2),

$$\begin{aligned} &|\mathrm{Tr}\left\{(S_{\Delta t}\sigma(X_k))(S_{\Delta t}\sigma(X_k))^* \mathcal{A}(s,t)\right\}| \\ &\le \mathrm{Tr}\left((-A)^{-1/2-\varepsilon}\right) \left|(-A)^{1/2+\varepsilon} S_{\Delta t}\right|_{\mathcal{L}(H)} |\sigma(X_k)|^2_{\mathcal{L}(H)} |\mathcal{A}(s,t)|_{\mathcal{L}(H)} \\ &\le c\Delta t^{-1/2-\varepsilon}(1+|X_k|)^2. \end{aligned}$$

We deduce by Lemma 4.1

$$b_k^{2,1} \le c\Delta t^{3/2-\varepsilon}. \tag{3.19}$$

[1] Recall that we in fact work with Galerkin approximations so that all sums below are finite sums.

The second term $b_k^{2,2}$ involves the same difficulty as a_k^1 above. We rewrite it using (2.11). This gives

$$\begin{aligned} b_k^{2,2} &= \mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\sum_{i\in\mathbb{N}}\left(A_{\Delta t}S_{\Delta t}^k x + A_{\Delta t}\Delta t\sum_{\ell=0}^{k-1}S_{\Delta t}^{k-\ell}f(X_\ell), Df_i(\tilde X(s))\right)\partial_i u(T-t,\tilde X(t))dsdt \\ &+\mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\sum_{i\in\mathbb{N}}\left(A_{\Delta t}\int_0^{t_k}S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})dW(\tau), Df_i(\tilde X(s))\right)\partial_i u(T-t,\tilde X(t))dsdt\end{aligned}$$

where, as above, $\ell_\tau=[\tau/\Delta t]$. The first term is bounded as follows, using Lemma 4.4, (2.3), (2.12), (2.14),

$$\begin{aligned} &\mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\sum_{i\in\mathbb{N}}\left(A_{\Delta t}S_{\Delta t}^k x + A_{\Delta t}\Delta t\sum_{\ell=0}^{k-1}S_{\Delta t}^{k-\ell}f(X_\ell), Df_i(\tilde X(s))\right)\partial_i u(T-t,\tilde X(t))dsdt \\ &= \mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\left(Df(\tilde X(s))\cdot\left(A_{\Delta t}S_{\Delta t}^k x + A_{\Delta t}\Delta t\sum_{\ell=0}^{k-1}S_{\Delta t}^{k-\ell}f(X_\ell)\right), Du(T-t,\tilde X(t))\right)dsdt \\ &\le c\mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}|(-A)^{\varepsilon}S_{\Delta t}|_{\mathcal{L}(H)}\left(\left|(-A)^{1-\varepsilon}S_{\Delta t}^k x\right| + \sum_{\ell=0}^{k-1}\left|(-A)^{1-\varepsilon}S_{\Delta t}^{k-\ell}\right|_{\mathcal{L}(H)}|f(X_\ell)|\right)ds\,dt \\ &\le c\Delta t^{2-\varepsilon}(t_k^{-1+\varepsilon}+1).\end{aligned}$$

The second term of $b_k^{2,2}$ requires an integration by parts, we obtain

$$\begin{aligned} &\mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\sum_{i\in\mathbb{N}}\left(A_{\Delta t}\int_0^{t_k}S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})dW(\tau), Df_i(\tilde X(s))\right)\partial_i u(T-t,\tilde X(t))dsdt \\ &= \mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\sum_{i,j,m\in\mathbb{N}}\left(A_{\Delta t}\int_0^{t_k}S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})e_m, e_j\right)d\beta_m(\tau)\partial_j f_i(\tilde X(s))\partial_i u(T-t,\tilde X(t))dsdt \\ &= \mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\int_0^{t_k}\sum_{i,j,m,n\in\mathbb{N}}\left(A_{\Delta t}S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})e_m, e_j\right)\Big[\partial_{j,n}f_i(\tilde X(s))\left(D_\tau^m\tilde X(s), e_n\right)\partial_i u(T-t,\tilde X(t)) \\ &\qquad +\partial_j f_i(\tilde X(s))\partial_{i,n}u(T-t,\tilde X(t))\left(D_\tau^m\tilde X(t), e_n\right)\Big]d\tau dsdt \\ &= \mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\int_0^{t_k}\sum_{i,m\in\mathbb{N}}D^2 f_i(\tilde X(s))\left(A_{\Delta t}S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})e_m, D_\tau^m\tilde X(s)\right)\partial_i u(T-t,\tilde X(t)) \\ &\qquad +\left(B_i(s,t)A_{\Delta t}S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})e_m, D_\tau^m\tilde X(t)\right)d\tau dsdt \\ &= \mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\int_0^{t_k}\sum_{i\in\mathbb{N}}\mathrm{Tr}\left\{\left(D_\tau\tilde X(s)\right)^* D^2 f_i(\tilde X(s))A_{\Delta t}S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})\right\}\partial_i u(T-t,\tilde X(t)) \\ &\qquad +\mathrm{Tr}\left\{\left(D_\tau\tilde X(t)\right)^* B_i(s,t)A_{\Delta t}S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})\right\}d\tau dsdt\end{aligned}$$

where, for $i \in \mathbb{N}$, $B_i(s,t)$ is defined by

$$(B_i(s,t)g,h) = (Df_i(\tilde{X}(s)),g)\sum_{n\in\mathbb{N}} \partial_{i,n}u(T-t,\tilde{X}(t))(h,e_n), \quad g,h\in H.$$

The first term above is estimate as $b_k^{2,1}$. For the second term, we write

$$\begin{aligned}\textstyle\sum_{i\in\mathbb{N}} (B_i(s,t)g,h) &= D^2u(T-t,\tilde{X}(t))\cdot(Df(\tilde{X}(s))\cdot g,h)\\ &= (D^2u(T-t,\tilde{X}(t))h, Df(\tilde{X}(s))\cdot g), \quad g,h\in H.\end{aligned}$$

Therefore

$$\left|\sum_{i\in\mathbb{N}} B_i(s,t)\right|_{\mathcal{L}(H)} \le \left|Df(\tilde{X}(s))\right|_{\mathcal{L}(H)}\left|D^2u(T-t,\tilde{X}(t))\right|_{\mathcal{L}(H)}.$$

We deduce by Lemma 4.3, (3.11), (2.3), (2.14), (2.2), (2.12), (2.2), Lemma 4.1 and similar arguments as above

$$\begin{aligned}&\mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\sum_{i\in\mathbb{N}}\left(A_{\Delta t}\int_0^{t_k} S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})dW(\tau), Df_i(\tilde{X}(s))\right)\partial_i u(T-t,\tilde{X}(t))dsdt\\ &\le c\Delta t^{3/2-\varepsilon}.\end{aligned}$$

Therefore

$$b_k^{2,2} \le c\Delta t^{3/2-\varepsilon}.$$

It is also easy to see that

$$\begin{aligned}b_k^{2,3} &= \mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\sum_{i\in\mathbb{N}}\left(S_{\Delta t}f(X_k), Df_i(\tilde{X}(s))\right)\partial_i u(T-t,\tilde{X}(t))dsdt\\ &= \mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t} Du(T-t,\tilde{X}(t))\cdot\left(Df(\tilde{X}(s))\cdot S_{\Delta t}f(X_k)\right)dsdt\\ &\le c\Delta t^2.\end{aligned}$$

It remains to estimate $b_k^{2,4}$. We again integrate by parts the stochastic integral and obtain by Lemma 2.1:

$$\begin{aligned}b_k^{2,4} &= \mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\sum_{i\in\mathbb{N}}\left(Df_i(\tilde{X}(s)),\sigma(X_k)dW(s)\right)\partial_i u(T-t,\tilde{X}(t))dt\\ &= \mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\mathrm{Tr}\left\{\left(D_s\tilde{X}(t)\right)^* D^2u(T-t,\tilde{X}(t))Df(\tilde{X}(s))\sigma(X_k)\right\}ds\,dt\\ &= \mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\mathrm{Tr}\left\{\sigma^*(X_k)S_{\Delta t}D^2u(T-t,\tilde{X}(t))Df(\tilde{X}(s))\sigma(X_k)\right\}ds\,dt\\ &\le c\Delta t^{3/2-\varepsilon},\end{aligned}$$

thanks to (3.16), (2.2) and (2.14).

We conclude this step by gathering the previous estimates. This enables us to write

$$\sum_{k=1}^{N-1} |b_k| \le c\Delta t^{1/2-\varepsilon}$$

Step 6: Estimate of c_k.
Using the symmetry of Du, we introduce the decomposition of c_k:

$$\begin{aligned}
c_k \quad &= \frac{1}{2}\mathbb{E}\int_{t_k}^{t_{k+1}} \mathrm{Tr}\left\{\left[\sigma(\tilde{X}(t))\sigma^*(\tilde{X}(t)) - (S_{\Delta t}\sigma(X_k))\,(S_{\Delta t}\sigma(X_k))^*\right] D^2u(T-t,\tilde{X}(t))\right\} dt \\
&= \frac{1}{2}\mathbb{E}\int_{t_k}^{t_{k+1}} \mathrm{Tr}\left\{(I-S_{\Delta t})\sigma(\tilde{X}(t))\left((I-S_{\Delta t})\sigma(\tilde{X}(t))\right)^* D^2u(T-t,\tilde{X}(t))\right\} dt \\
&+\mathbb{E}\int_{t_k}^{t_{k+1}} \mathrm{Tr}\left\{S_{\Delta t}\sigma(\tilde{X}(t))\left((I-S_{\Delta t})\sigma(\tilde{X}(t))\right)^* D^2u(T-t,\tilde{X}(t))\right\} dt \\
&+\frac{1}{2}\mathbb{E}\int_{t_k}^{t_{k+1}} \mathrm{Tr}\left\{S_{\Delta t}\left(\sigma(\tilde{X}(t)) - \sigma(X_k)\right)\left(S_{\Delta t}\sigma(\tilde{X}(t))\right)^* D^2u(T-t,\tilde{X}(t))\right\} dt \\
&+\frac{1}{2}\mathbb{E}\int_{t_k}^{t_{k+1}} \mathrm{Tr}\left\{S_{\Delta t}\sigma(X_k)\left(S_{\Delta t}\sigma(\tilde{X}(t)) - \sigma(X_k)\right)^* D^2u(T-t,\tilde{X}(t))\right\} dt \\
&= c_k^1 + c_k^2 + c_k^3 + c_k^4.
\end{aligned}$$

The first two terms are easy to treat, we use similar arguments as in the previous steps and write thanks to (2.7), Lemma 4.5, Lemma 4.1, (2.15)

$$\begin{aligned}
c_k^1 \quad &\le c\mathbb{E}\int_{t_k}^{t_{k+1}} \mathrm{Tr}\left\{(-A)^{-1/2+\varepsilon}(I-S_{\Delta t})\sigma(\tilde{X}(t)\sigma^*(\tilde{X}(t))(I-S_{\Delta t})(-A)^{-1/2+\varepsilon}\right\}(T-t)^{-1+2\varepsilon}dt \\
&\le c\mathbb{E}\int_{t_k}^{t_{k+1}} \mathrm{Tr}\left\{(-A)^{-1/2+\varepsilon}(I-S_{\Delta t})(I-S_{\Delta t})(-A)^{-1/2+\varepsilon}\right\}(T-t)^{-1+2\varepsilon}dt \\
&\le c\Delta t^{1/2-3\varepsilon}\int_{t_k}^{t_{k+1}} (T-t)^{-1+\varepsilon}dt
\end{aligned}$$

The second term is similar, we have

$$\begin{aligned}
c_k^2 \quad &\le \mathbb{E}\int_{t_k}^{t_{k+1}} \left|(-A)^{-1/2+\varepsilon}(I-S_{\Delta t})\right|_{\mathcal{L}(H)} \left|\sigma(\tilde{X}(t)\right|^2_{\mathcal{L}(H)} \left|(-A)^{2\varepsilon}S_{\Delta t}\right|_{\mathcal{L}(H)} \mathrm{Tr}\left\{(-A)^{-1/2-\varepsilon}\right\} \\
&\qquad\qquad \left|(-A)^{1/2-\varepsilon}D^2u(T-t,\tilde{X}(t))(-A)^{1/2-\varepsilon}\right|_{\mathcal{L}(H)} dt \\
&\le c\Delta t^{1/2-3\varepsilon}\int_{t_k}^{t_{k+1}} (T-t)^{-1+\varepsilon}dt
\end{aligned}$$

The estimate of the next term is much more complicated. It is based on similar arguments as before but the computations are much longer.

We use (3.3) and obtain for $h, k \in H$:

$$\begin{aligned}
\left(\left(\sigma(\tilde{X}(t)) - \sigma(X_k)\right) h, k\right) &= \frac{1}{2}\int_{t_k}^{t} \mathrm{Tr}\left\{\left(S_{\Delta t}\sigma(X_k)\right)\left(S_{\Delta t}\sigma(X_k)\right)^* D^2\left(\sigma(\cdot)h,k\right)(\tilde{X}(s))\right\} dt \\
&+\frac{1}{2}\int_{t_k}^{t}\left(A_{\Delta t}X_k + S_{\Delta t}f(X_k), D\left(\sigma(\cdot)h,k\right)(\tilde{X}(s))\right) dt \\
&= (\mathcal{A}h,k) + (\mathcal{B}h,k) + (\mathcal{C}h,k).
\end{aligned}$$

Thus we may write

$$\begin{aligned}
c_k^3 &= \frac{1}{2}\mathbb{E}\int_{t_k}^{t_{k+1}} \mathrm{Tr}\left\{S_{\Delta t}\mathcal{A}\left(S_{\Delta t}\sigma(\tilde{X}(t))\right)^* D^2u(T-t,\tilde{X}(t))\right\} dt \\
&+\frac{1}{2}\mathbb{E}\int_{t_k}^{t_{k+1}} \mathrm{Tr}\left\{S_{\Delta t}\mathcal{B}\left(S_{\Delta t}\sigma(\tilde{X}(t))\right)^* D^2u(T-t,\tilde{X}(t))\right\} dt \\
&+\frac{1}{2}\mathbb{E}\int_{t_k}^{t_{k+1}} \mathrm{Tr}\left\{S_{\Delta t}\mathcal{C}\left(S_{\Delta t}\sigma(\tilde{X}(t))\right)^* D^2u(T-t,\tilde{X}(t))\right\} dt \\
&= c_k^{3,1} + c_k^{3,2} + c_k^{3,3}.
\end{aligned}$$

Note that

$$(\mathcal{A}h,k) = \frac{1}{2}\int_{t_k}^{t}\sum_{\ell\in\mathbb{N}}\left(\left(\sigma''(\tilde{X}(s)).(S_{\Delta t}\sigma(X_k)e_\ell, S_{\Delta t}\sigma(X_k)e_\ell)\right) h, k\right) ds.$$

By (2.5), for $u, v \in H$,

$$\left(\left(\sigma''(\tilde{X}(s)).(u,v)\right) h, k\right) \le L_\sigma |u|_{-1/4}|v|_{-1/4}\,|h|\,|k| \le c|u|\,|v|\,|h|\,|k|.$$

We deduce, thanks to (2.2), (2.14),

$$(\mathcal{A}h,k) \le c\Delta t^{1/2-\varepsilon}(1+|X_k|)^2|h|\,|k|,$$

and

$$|\mathcal{A}|_{\mathcal{L}(H)} \le c\Delta t^{1/2-\varepsilon}(1+|X_k|)^2.$$

Then, by Lemma 4.1, Lemma 4.5, (2.14) and again (2.2)

$$c_k^{3,1} \le c\Delta t^{1/2-3\varepsilon}\int_{t_k}^{t_{k+1}}(T-t)^{-1/2+\varepsilon}dt.$$

The term $c_k^{3,2}$ involves the same difficulty as a_k and $b_k^{2,2}$. We use (2.11) to replace X_k by a sum of three terms:

$$\begin{aligned}
(\mathcal{B}h,k) &= \frac{1}{2}\int_{t_k}^{t}\Big(A_{\Delta t}S_{\Delta t}^k x+\Delta t A_{\Delta t}\sum_{\ell=0}^{k-1}S_{\Delta t}^{k-\ell}f(X_\ell)+\int_0^{t_k}A_{\Delta t}S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})dW(\tau),\\
&\qquad\qquad D\left(\sigma(\cdot)h,k\right)(\tilde X(s))\Big)ds\\
&= \frac{1}{2}\int_{t_k}^{t}\Big(\Big[\sigma'(\tilde X(s))\cdot\big(A_{\Delta t}S_{\Delta t}^k x+\Delta t A_{\Delta t}\sum_{\ell=0}^{k-1}S_{\Delta t}^{k-\ell}f(X_\ell)\\
&\qquad\qquad +\textstyle\int_0^{t_k}A_{\Delta t}S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})dW(\tau)\big)\Big]h,k\Big)ds\\
&= (\mathcal{B}_1h,k)+(\mathcal{B}_2h,k)+(\mathcal{B}_3h,k).
\end{aligned}$$

We then write thanks to (2.4), (2.12) and (2.14)

$$\begin{aligned}
(\mathcal{B}_1h,k) &= \frac{1}{2}\int_{t_k}^{t}\Big(\Big[\sigma'(\tilde X(s))\cdot A_{\Delta t}S_{\Delta t}^k x\Big]h,k\Big)ds\\
&\le c\int_{t_k}^{t}\Big|\sigma'(\tilde X(s))\cdot A_{\Delta t}S_{\Delta t}^k x\Big|_{\mathcal{L}(H)}|h|\,|k|ds\\
&\le c\Delta t|A_{\Delta t}S_{\Delta t}^k x|\,|h|\,|k|\\
&\le c\Delta t^{1-\varepsilon}t_k^{1-\varepsilon}\,|h|\,|k|.
\end{aligned}$$

Similarly

$$(\mathcal{B}_2h,k)\le c\Delta t^{1-\varepsilon}\,|h|\,|k|.$$

It follows, thanks to Lemma 4.5, (2.14) and (2.2)

$$\begin{aligned}
&\frac{1}{2}\mathbb{E}\int_{t_k}^{t_{k+1}}\mathrm{Tr}\Big\{S_{\Delta t}(\mathcal{B}_1+\mathcal{B}_2)\Big(S_{\Delta t}\sigma(\tilde X(t))\Big)^*D^2u(T-t,\tilde X(t))\Big\}dt\\
&\le c\Delta t^{1-3\varepsilon}(t_k^{1-\varepsilon}+1)\int_{t_k}^{t_{k+1}}(T-t)^{-1/2+\varepsilon}dt.
\end{aligned}$$

The estimate of the part of $c_k^{3,2}$ involving $\mathcal{B}_3$ is very technical. As before, we get rid of the stochastic integral thanks to an integration by parts. This results in a supplementary trace term. In order to work with the double trace, we write everything in terms of the components of the operators and vectors. Given an operator G on H, we set $G^{i,j}=(Ge_i,e_j)$. We thus write

$$
\begin{aligned}
&\mathbb{E}\int_{t_k}^{t_{k+1}} \mathrm{Tr}\left\{S_{\Delta t}\mathcal{B}_3\left(S_{\Delta t}\sigma(\tilde{X}(t))\right)^* D^2u(T-t,\tilde{X}(t))\right\}dt \\
&= \mathbb{E}\int_{t_k}^{t_{k+1}} \sum_{i,j,m\in\mathbb{N}} \mathcal{B}_3^{i,j}\sigma^{m,j}(\tilde{X}(t)))\left(S_{\Delta t}D^2u(T-t,\tilde{X}(t))S_{\Delta t}\right)^{m,i}dt \\
&= \sum_{i,j,m,n,r\in\mathbb{N}} \mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\int_0^{t_k} \partial_r\sigma^{i,j}(\tilde{X}(s))\left(A_{\Delta t}S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})e_n,e_r\right)d\beta_n(\tau)\sigma^{m,j}(\tilde{X}(t))) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\left(S_{\Delta t}D^2u(T-t,\tilde{X}(t))S_{\Delta t}\right)^{m,i}ds\,dt.
\end{aligned}
$$

It is important to recall here that in fact we work with finite dimensional approximations of the solutions so that all the above sums are finite. We now use the Malliavin integration by parts and obtain

$$
\begin{aligned}
&\mathbb{E}\int_{t_k}^{t_{k+1}} \mathrm{Tr}\left\{S_{\Delta t}\mathcal{B}_3\left(S_{\Delta t}\sigma(\tilde{X}(t))\right)^* D^2u(T-t,\tilde{X}(t))\right\}dt \\
&= \sum_{i,j,m,n,r\in\mathbb{N}} \mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\int_0^{t_k} \sum_{p\in\mathbb{N}}\partial_{r,p}\sigma^{i,j}(\tilde{X}(s))\left(D_\tau^n\tilde{X}(s),e_p\right)\left(A_{\Delta t}S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})e_n,e_r\right)\sigma^{m,j}(\tilde{X}(t))) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\left(S_{\Delta t}D^2u(T-t,\tilde{X}(t))S_{\Delta t}\right)^{m,i} \\
&+\sum_{p\in\mathbb{N}}\partial_r\sigma^{i,j}(\tilde{X}(s))\left(A_{\Delta t}S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})e_n,e_r\right)\partial_p\sigma^{m,j}(\tilde{X}(t)))\left(D_\tau^n\tilde{X}(t),e_p\right)\left(S_{\Delta t}D^2u(T-t,\tilde{X}(t))S_{\Delta t}\right)^{m,i} \\
&+\partial_r\sigma^{i,j}(\tilde{X}(s))\left(A_{\Delta t}S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})e_n,e_r\right)\sigma^{m,j}(\tilde{X}(t)))\left(S_{\Delta t}\left(D^3u(T-t,\tilde{X}(t))\cdot D_\tau^n\tilde{X}(s)\right)S_{\Delta t}\right)^{m,i}d\tau\,ds\,dt \\
&= I+II+III.
\end{aligned}
$$

We then write

$$
\begin{aligned}
I &= \sum_{i,j,m,n\in\mathbb{N}} \mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\int_0^{t_k} D^2\sigma^{i,j}(\tilde{X}(s))\cdot\left(D_\tau^n\tilde{X}(s),A_{\Delta t}S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})e_n\right)\sigma^{m,j}(\tilde{X}(t)) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\left(S_{\Delta t}D^2u(T-t,\tilde{X}(t))S_{\Delta t}\right)^{m,i}d\tau\,ds\,dt \\
&= \sum_{j\in\mathbb{N}}\mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\int_0^{t_k} D^2u(T-t,\tilde{X}(t))\cdot\left(\phi_1(s,\tau,k)e_j,S_{\Delta t}\sigma(\tilde{X}(t))e_j\right)d\tau\,ds\,dt \\
&= \mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\int_0^{t_k}\mathrm{Tr}\left\{\sigma^*(\tilde{X}(t))S_{\Delta t}D^2u(T-t,\tilde{X}(t))\phi_1(s,\tau,k)\right\}d\tau\,ds\,dt
\end{aligned}
$$

where we have set

$$\phi_1(s,\tau,k)h_1 = \sum_{n\in\mathbb{N}} S_{\Delta t}\left(D^2\sigma(\tilde X(s))\cdot\left(D^n_\tau \tilde X(s), A_{\Delta t}S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})e_n\right)\right)h_1,\ h_1\in H.$$

Let us define Σ_{s,h_1,h_2} by

$$(\Sigma_{s,h_1,h_2}u,v) = \left(S_{\Delta t}\left(D^2\sigma(\tilde X(s))\cdot(u,v)\right)h_1,h_2\right),\ u,v\in H.$$

Then by (2.5)

$$|\Sigma_{s,h_1,h_2}|_{\mathcal{L}(H)} \le c\,|h_1|\,|h_2|.$$

We deduce by (2.4), (2.2), (2.14), (2.12), (3.11) and Lemma 4.3

$$\begin{aligned}(\phi_1(s,\tau,k)h_1,h_2) &= \mathrm{Tr}\left\{\sigma^*(X_{\ell_\tau})S_{\Delta t}^{k-\ell_\tau}A_{\Delta t}\Sigma_{s,h_1,h_2}D_\tau\tilde X(s)\right\}\\ &\le |\sigma^*(X_{\ell_\tau})|_{\mathcal{L}(H)}\mathrm{Tr}(-A)^{-1/2-\varepsilon}|(-A)^{1/2+\varepsilon}S_{\Delta t}^{k-\ell_\tau}A_{\Delta t}|_{\mathcal{L}(H)}|\Sigma_{s,h_1,h_2}|_{\mathcal{L}(H)}|D_\tau\tilde X(s)|_{\mathcal{L}(H)}\\ &\le c\Delta t^{-1/2-2\varepsilon}t_k^{-1+\varepsilon}|h_1|\,|h_2|(1+|X_k|)\end{aligned}$$

and by Lemma 4.1, Lemma 4.5 and (2.2)

$$I \le c\Delta t^{1/2-2\varepsilon}t_k^{-1+\varepsilon}\int_{t_k}^{t_{k+1}}(T-t)^{-1/2-\varepsilon}dt.$$

Similarly, we may write

$$\begin{aligned}II = \sum_{n\in\mathbb{N}}\mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\int_0^{t_k}\mathrm{Tr}\Big\{&\left[\left(D\sigma(\tilde X(s))\cdot\left(A_{\Delta t}S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})e_n\right)\right)\right]\\ &\left[\left(D\sigma(\tilde X(t))\cdot\left(D^n_\tau\tilde X(t)\right)\right)\right]^* S_{\Delta t}D^2u(T-t,\tilde X(t))S_{\Delta t}\Big\}d\tau\,ds\,dt\\ \le c\Delta t^{-2\varepsilon}\mathbb{E}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\int_0^{t_k}&|\phi_2(\tau,s,t,k)|_{\mathcal{L}(H)}\,(T-t)^{-1/2+\varepsilon}d\tau\,ds\,dt\end{aligned}$$

with

$$\phi_2(\tau,s,t,k) = \sum_{n\in\mathbb{N}}\left[\left(D\sigma(\tilde X(s))\cdot\left(A_{\Delta t}S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})e_n\right)\right)\right]\left[\left(D\sigma(\tilde X(t))\cdot\left(D^n_\tau\tilde X(t)\right)\right)\right]^*.$$

We use similar arguments to estimate its norm. For $u,v\in H$, we write

$$\begin{aligned}&(\phi_2(\tau,s,t,k)u,v)\\ &= \textstyle\sum_{n\in\mathbb{N}}\left(\left[\left(D\sigma(\tilde X(t))\cdot\left(D^n_\tau\tilde X(t)\right)\right)\right]^*u, \left[\left(D\sigma(\tilde X(s))\cdot\left(A_{\Delta t}S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})e_n\right)\right)\right]^*v\right)\\ &= \mathrm{Tr}\left\{\sigma^*(X_{\ell_\tau})S_{\Delta t}^{k-\ell_\tau}A_{\Delta t}a_v^*b_u\right\}\end{aligned}$$

with

$$a_vh = \left[S_{\Delta t}\left(D\sigma(\tilde X(s))\cdot h\right)\right]^*v,$$

$$b_u h = \left[S_{\Delta t}\left(D\sigma(\tilde{X}(t))\cdot\left(D_\tau^h\tilde{X}(t)\right)\right)\right]^* u.$$

Since

$$|a_v|_{\mathcal{L}(H)} \le c\,|v|, \quad |b_u|_{\mathcal{L}(H)} \le c\,|u|,$$

we deduce

$$|\phi_2(\tau,s,t,k)|_{\mathcal{L}(H)} \le c\mathrm{Tr}\{S_{\Delta t}^{k-\ell_\tau}A_{\Delta t}\} \le c\Delta t^{-1/2-2\varepsilon}t_{k-\ell_\tau}^{-1+\varepsilon}.$$

and

$$II \le c\Delta t^{1/2-4\varepsilon}\int_{t_k}^{t_{k+1}}(T-t)^{-1/2-\varepsilon}dt.$$

Finally

$$III = \frac{1}{2}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\int_0^{t_k}\sum_{n\in\mathbb{N}}\mathrm{Tr}\{\gamma_n S_{\Delta t}\sigma(\tilde{X}(t))\}d\tau\,ds\,dt$$

where for $u,v\in H$

$$\begin{aligned}&\sum_{n\in\mathbb{N}}(\gamma_n u,v)\\ &=\sum_{n\in\mathbb{N}}D^3u(T-t,\tilde{X}(t))\left(D_\tau^n\tilde{X}(s),u,S_{\Delta t}\left(D\sigma(\tilde{X}(s))\cdot\left(A_{\Delta t}S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})e_n\right)\right)v\right)\\ &=\mathrm{Tr}\left\{\kappa(u,v)(-A)^{-1/2-\varepsilon}(-A)^{2\varepsilon}D_\tau^n\tilde{X}(s)\right\},\end{aligned}$$

and for $h_1,h_2\in H$

$$(\kappa(u,v)h_1,h_2) = D^3u(T-t,\tilde{X}(t))\cdot\left((-A)^{1/2-\varepsilon}h_1,u,S_{\Delta t}\left(D\sigma(\tilde{X}(s))\cdot\left(A_{\Delta t}S_{\Delta t}^{k-\ell_\tau}\sigma(X_{\ell_\tau})h_2\right)\right)v\right).$$

By Lemma 4.6

$$|\kappa(u,v)|_{\mathcal{L}(H)} \le c(T-t)^{-1/2+\varepsilon}t_{k-\ell_\tau}^{-1+3\varepsilon}\Delta t^{-3\varepsilon}|u|\,|v|.$$

Therefore, by (2.2), (3.11) and Lemma 4.3,

$$\sum_{n\in\mathbb{N}}(\gamma_n u,v) \le c(T-t)^{-1/2+\varepsilon}t_{k-\ell_\tau}^{-1+\varepsilon}\Delta t^{-3\varepsilon}|u|\,|v|.$$

It follows

$$|\gamma_n|_{\mathcal{L}(H)} \le c(T-t)^{-1/2+\varepsilon}t_{k-\ell_\tau}^{-1+\varepsilon}\Delta t^{-3\varepsilon}$$

and by (2.2), (2.14)

$$III \le c\Delta t^{1/2-4\varepsilon}\int_{t_k}^{t_{k+1}}(T-t)^{-1/2+\varepsilon}dt.$$

We can now conclude

$$|c_k^{3,2}| \le c\Delta t^{1/2-4\varepsilon}(t_k^{-1+\varepsilon}+1)(\int_{t_k}^{t_{k+1}}(T-t)^{-1/2+\varepsilon}dt+1).$$

Finally, it is easy to check

$$|\mathcal{C}|_{\mathcal{L}(H)} \le c\Delta t(1+|X_k|)$$

and

$$|c_k^{3,3}| \le c\Delta t^{1-2\varepsilon}\int_{t_k}^{t_{k+1}}(T-t)^{-1/2+\varepsilon}dt.$$

We deduce

$$|c_k^3| \leq c\Delta t^{1/2-4\varepsilon}(t_k^{-1+\varepsilon}+1)\int_{t_k}^{t_{k+1}}((T-t)^{-1/2-\varepsilon}+1)dt,$$

and, since c_k^4 is majorized in exactly the same way,

$$|c_k| \leq c\Delta t^{1/2-4\varepsilon}(t_k^{-1+\varepsilon}+1)\int_{t_k}^{t_{k+1}}((T-t)^{-1+\varepsilon}+1)dt.$$

It follows

$$\sum_{k=1}^{N-1}|c_k| \leq c\Delta t^{1/2-4\varepsilon}.$$

Step 7: Conclusion.

It is now easy to gather all previous estimates in (3.7) and deduce

$$|u(T,x)-\mathbb{E}\left(\varphi(X_N)\right)| \leq c\Delta t^{1/2-4\varepsilon}.$$

Recall that all the above computations have been done on the Galerkin approximations of X and X_k). The constant c above does not depend on m so that we can easily let $m\to\infty$ in this estimate and obtain the result.

4. Auxiliary Lemmas

In this section, we state and prove technical Lemmas used in the preceeding section. Again, the various estimates used here could be difficult to justify rigorously on the infinite dimensional equation and we in fact work with Galerkin approximations. Taking the limit $m\to\infty$ at the end of the proofs gives the results rigorously.

The first two Lemmas are very classical and we state them without proof.

Lemma 4.1. *For any $l\in\mathbb{N}$, there exists a constant c_l such that*

$$\max_{k=0,\dots,N}\mathbb{E}(|X_k|^l) \leq c_l(|x|^l+1).$$

Lemma 4.2. *For any $l\in\mathbb{N}$, there exists a constant $\tilde{c}_l$ such that*

$$\sup_t \mathbb{E}(|X(t,x)|) \leq \tilde{c}_l(|x|^l+1).$$

Lemma 4.3. *For any $\beta\in[0,1/4)$, there exists a constant c such for $k=1,\dots,N$, $s\in[0,t_k]$, we have*

$$t_{k-\ell_s}^{2\beta}\mathbb{E}\left(|D_s^hX_k|_\beta^2\right) \leq c|h|^2,\ h\in H.$$

Proof: By (2.11) and the chain rule, we obtain the following formula for the Malliavin derivative of X_k:

$$\begin{aligned} D_s^hX_k &= S_{\Delta t}^{k-\ell_s}\sigma(X_{\ell_s})h + \Delta t\sum_{\ell=\ell_s+1}^{k-1}S_{\Delta t}^{k-\ell}f'(X_\ell)\cdot D_s^hX_\ell \\ &+\sqrt{\Delta t}\sum_{\ell=\ell_s+1}^{k-1}S_{\Delta t}^{k-\ell}(\sigma'(X_\ell)\cdot D_s^hX_\ell)\chi_{\ell+1} \end{aligned}$$

for $s \in [0, t_k]$ and $h \in H$.

By (2.12), (2.2),(2.3), (2.4), we deduce for $\varepsilon > 0$

$$\begin{aligned}
\mathbb{E}\left(|D_s^h X_k|_\beta^2\right) &\leq c\Bigg(t_{k-\ell_s}^{-2\beta}|h|^2 + \left(\Delta t \sum_{\ell=\ell_s+1}^{k-1} \left|(-A)^\beta S_{\Delta t}^{k-\ell}\right|_{\mathcal{L}(H)} \left|f'(X_\ell)\right|_{\mathcal{L}(H)} \left|D_s^h X_\ell\right|\right)^2 \\
&\quad + \Delta t \sum_{\ell=\ell_s+1}^{k-1} \left|(-A)^\beta S_{\Delta t}^{k-\ell}(\sigma'(X_\ell)\cdot D_s^h X_\ell)\right|_{\mathcal{L}_2(H)}^2 \Bigg) \\
&\leq c\Bigg(t_{k-\ell_s}^{-2\beta}|h|^2 + L_F^2 \left(\Delta t \sum_{\ell=\ell_s+1}^{k-1} t_{k-\ell}^{-\beta} \left|D_s^h X_\ell\right|\right)^2 \\
&\quad + L_\sigma \Delta t \sum_{\ell=\ell_s+1}^{k-1} t_{k-\ell}^{-1/2-\varepsilon-2\beta} \left|D_s^h X_\ell)\right|_{\mathcal{L}_2(H)}^2 \Bigg).
\end{aligned}$$

It is now easy to use a discrete Gronwall Lemma and prove

$$\max_{l=\ell_s+1,\dots,k} t_{\ell-\ell_s}^{2\beta} \mathbb{E}\left(|D_s^h X_\ell|^2\right) \leq c|h|^2$$

□

Lemma 4.4. *Let $\varphi \in C_b^1(H, \mathbb{R})$. For any $\beta < 1/2$, there exists a constant c_β such that for $t > 0$, $x \in H$*

$$|Du(t,x)|_\beta \leq c_\beta t^{-\beta} \|\varphi\|_1,$$

where u is defined in (3.4).

Proof: Differentiating (3.4), we obtain for $h \in H$:

$$Du(t,x)\cdot h = \mathbb{E}\left(D\varphi(X(t,x))\cdot \eta^{h,x}(t)\right)$$

where $\eta^{h,x}(t)$ is the solution of

$$\begin{cases} d\eta^{h,x} = \left(A\eta^{h,x} + f'(X(t,x))\cdot\eta^{h,x}\right)dt + \sigma'(X(t,x))\cdot\eta^{h,x}dW, \\ \eta^{h,x}(0) = h. \end{cases}$$

We rewrite this equation in the integral form

$$\eta^{h,x}(t) = S(t)h + \int_0^t S(t-s)f'(X(s,x))\cdot\eta^{h,x}(s)ds + \int_0^t S(t-s)\sigma'(X(s,x))\cdot\eta^{h,x}(s)dW(s),\ t\geq 0.$$

By (2.4), (2.2), (2.12), we have for $y, k \in H$ and $\alpha > 1/2$:

$$\left|S(t)\sigma'(y)\cdot k\right|_{\mathcal{L}_2(H)} \leq L_\sigma \left|(-A)^{-\alpha/2}\right|_{\mathcal{L}_2(H)} \left|(-A)^{\alpha/2}S(t)\right|_{\mathcal{L}(H)} |k| \leq ct^{-\alpha/2}|k|.$$

Using (2.3) and then Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}\mathbb{E}\left(\left|\eta^{h,x}(t)\right|^2\right) &\leq c\, t^{-2\beta}|h|^2_{-\beta} + L_f^2\mathbb{E}\left(\left(\int_0^t |\eta^{h,x}(s)|ds\right)^2\right) + \mathbb{E}\int_0^t (t-s)^{-\alpha}\left|\eta^{h,x}(s)\right|^2 ds\\ &\leq c\, t^{-2\beta}|h|^2_{-\beta} + c\int_0^t \mathbb{E}(|\eta^{h,x}(s)|^2 ds + \mathbb{E}\int_0^t (t-s)^{-\alpha}\left|\eta^{h,x}(s)\right|^2 ds.\end{aligned}$$

It is classical that this implies

$$\sup_{t\in[0,T]} t^{2\beta}\mathbb{E}\left(\left|\eta^{h,x}(t)\right|^2\right) \leq |h|^2_{-\beta}. \tag{4.1}$$

We deduce

$$|Du(t,x)\cdot h| \leq c\|\varphi\|_1 t^{-\beta}|h|_{-\beta}.$$

Taking the supremum over h yields the result.

□

Lemma 4.5. *Let $\varphi \in C_b^2(H,\mathbb{R})$. For any $\beta,\gamma < 1/2$, there exists a constant $c_{\beta,\gamma}$ such that for $t>0$, $x\in H$*

$$|(-A)^\beta D^2u(t,x)(-A)^\gamma|_{\mathcal{L}(H)} \leq c_{\beta,\gamma}t^{-(\beta+\gamma)}\|\varphi\|_2,$$

where u is defined in (3.4).

Proof: We use the same notations as in the proof of Lemma 4.4. We differentiate a second time (3.4) and obtain for $h,k\in H$:

$$D^2u(t,x)\cdot(h,k) = \mathbb{E}\left(D^2\varphi(X(t,x))\cdot(\eta^{h,x}(t),\eta^{k,x}(t)) + D\varphi(X(t,x))\cdot\zeta^{h,k,x}(t)\right) \tag{4.2}$$

where $\zeta^{h,k,x}(t)$ is the solution of

$$\left\{\begin{aligned} d\zeta^{h,k,x} &= \left(A\zeta^{h,k,x} + f''(X(t,x))\cdot(\eta^{h,x}(t),\eta^{k,x}(t)) + f'(X(t,x))\cdot\zeta^{h,k,x}(t)\right)dt\\ &\quad+\left(\sigma''(X(t,x))\cdot(\eta^{h,x}(t),\eta^{k,x}(t)) + \sigma'(X(t,x))\cdot\zeta^{h,k,x}(t)\right)dW,\\ \zeta^{h,k,x}(0) &= 0.\end{aligned}\right.$$

We rewrite this equation in the integral form

$$\begin{aligned}\zeta^{h,k,x}(t) &= \int_0^t S(t-s)\left(f''(X(s,x))\cdot(\eta^{h,x}(s),\eta^{k,x}(s)) + f'(X(s,x))\cdot\zeta^{h,k,x}(s)\right)ds\\ &\quad+\int_0^t S(t-s)\left(\sigma''(X(s,x))\cdot(\eta^{h,x}(s),\eta^{k,x}(s)) + \sigma'(X(s,x))\cdot\zeta^{h,k,x}(s)\right)dW(s),\ t\geq 0.\end{aligned}$$

Using similar argument as above and (2.5), we prove

$$\begin{aligned}\mathbb{E}\left(|\zeta^{h,k,x}(t)|^2\right) &\leq c\mathbb{E}\left(\int_0^t |\eta^{h,x}(s)|\,|\eta^{k,x}(s)| + |\zeta^{h,k,x}(s)|ds\right)^2\\ &\quad+c\mathbb{E}\int_0^t (t-s)^{-\alpha}\left(|\eta^{h,x}(s)|^2_{-1/4}|\eta^{k,x}(s)|^2_{-1/4} + |\zeta^{h,k,x}(s)|^2\right)ds,\ t\geq 0.\end{aligned} \tag{4.3}$$

Proceeding as in Lemma 4.4, we have thanks to Burkholder inequality and then to Minkowsky inequality

$$\begin{aligned}\mathbb{E}\left(\left|\eta^{h,x}(t)\right|^4\right) &\le c\, t^{-4\beta}|h|_{-\beta}^4 + c\mathbb{E}\left(\left(\int_0^t |\eta^{h,x}(s)|ds\right)^4\right) + \mathbb{E}\left(\left(\int_0^t (t-s)^{-\alpha}\left|\eta^{h,x}(s)\right|^2 ds\right)^2\right)\\ &\le c\, t^{-4\beta}|h|_{-\beta}^4 + c\left(\int_0^t \left(\mathbb{E}(|\eta^{h,x}(s)|^4)\right)^{1/4} ds\right)^4\\ &+c\left(\int_0^t (t-s)^{-\alpha}\left(\mathbb{E}(|\eta^{h,x}(s)|^4)\right)^{1/2} ds\right)^2.\end{aligned}$$

Taking the square root of this inequality and using a generalized Gronwall Lemma, we deduce

$$\sup_{t\in[0,T]} t^{4\beta}\mathbb{E}\left(\left|\eta^{h,x}(t)\right|^4\right) \le c|h|_{-\beta}^4. \tag{4.4}$$

Similarly, we have

$$\begin{aligned}\mathbb{E}(|\eta^{h,x}(t)|_{-1/4}^4) &\le ct^{1-4\beta}|h|_{-\beta}^4 + c\mathbb{E}\left(\int_0^t |\eta^{h,x}(s)|ds\right)^4 + c\mathbb{E}\left(\int_0^t (t-s)^{-\alpha}|\eta^{h,x}(s)|^2 ds\right)^2\\ &\le ct^{1-4\beta}|h|_{-\beta}^4 + c\left(\int_0^t \mathbb{E}\left(|\eta^{h,x}(s)|^4\right)^{1/4} ds\right)^4\\ &+c\left(\int_0^t (t-s)^{-\alpha}\mathbb{E}\left(|\eta^{h,x}(s)|^4\right)^{1/2} ds\right)^2\end{aligned}$$

Therefore, by (4.4),

$$\mathbb{E}(|\eta^{h,x}(t)|_{-1/4}^4) \quad \le ct^{1-4\beta}|h|_{-\beta}^4 \tag{4.5}$$

Plugging these inequalities and similar ones for $\eta^{k,x}$ in (4.3) yields

$$\sup_{t\in[0,T]} \mathbb{E}\left(\left|\zeta^{h,x}(t)\right|^2\right) \le c|h|_{-\beta}^2|h|_{-\gamma}^2.$$

The result follows easily using (4.4) and this inequality in (4.2).

□

The following Lemma is proved thanks to similar arguments.

Lemma 4.6. *Let $\varphi \in C_b^3(H,\mathbb{R})$. For any $\beta < 1/2$, there exists a constant c_β such that for $t > 0$, $x \in H$, $h_1 \in D((-A)^\beta)$, $h_2 \in H$, $h_3 \in H$*

$$D^3u(t,x)\cdot((-A)^\beta h_1, h_2, h_3) \le c_\beta t^{-\beta}\|\varphi\|_3|h_1|\,|h_2|\,|h_3|,$$

where u is defined in (3.4).

References

[1] E.J. Allen, S.J. Novosel, Z. Zhang, *Finite element and difference approximation of some linear stochastic partial differential equations*, Stochastics Stochastics Rep. **64** (1998), n^o 1-2, 117–142.

[2] V. Bally, D. Talay *The law of the Euler scheme for stochastic differential equations. I. Convergence rate of the distribution function*, Probab. Theory Related Fields 104, no. 1, 43–60, 1996.

[3] E. Buckwar, T. Shardlow, *Weak approximation of stochastic differential delay equations*, IMA J. Numer. Anal. 25, no. 1, 57–86, 2005.

[4] A. de Bouard, A. Debussche *A semi-discrete scheme for the stochastic nonlinear Schrödinger equation*, Numer. Math., 96, no 1, 2003.

[5] G. Da Prato, J. Zabczyk, *Stochastic equations in infinite dimensions*, in "Encyclopedia of Mathematics and its Application", Cambridge University Press, Cambridge, 1992.

[6] A.M. Davie, J.G. Gaines, *Convergence of numerical schemes for the solution of parabolic stochastic partial differential equations*, Math. Comp. **70** (2001), n^o 233, 121–134

[7] A. de Bouard, A. Debussche *Weak and strong order of convergence of a semi discrete scheme for the stochastic Nonlinear Schrodinger equation*, Applied Mathematics and Optimization Journal **54** (2006), n^o 3, 369–399.

[8] A. Debussche, J. Printems *Weak order for the discretization of the stochastic heat equation*, preprint.

[9] *Introduction to the theory of linear nonselfadjoint operators in Hilbert space* , Amer. Math. Soc., Providence, R. I., 1970.

[10] W. Greksch, P.E. Kloeden *Time-discretised Galerkin approximations of parabolic stochastic PDEs*, Bull. Austral. Math. Soc. **54** (1996), n^o 1, 79–85.

[11] I. Gyöngy, *Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise. I*, Potential Anal. 9, no. 1, 1–25, 1998.

[12] I. Gyöngy, *Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise. II*, Potential Anal. 11, no. 1, 1–37, 1999.

[13] I. Gyöngy, A. Millet, *On discretization schemes for stochastic evolution equations*, , Potential Analysis **23** (2005), n^o 2, 99–134.'

[14] I. Gyöngy, A. Millet, *Rate of Convergence of Implicit Approximations for stochastic evolution equations*, Stochastic Differential Equations: Theory and Applications. A volume in Honor of Professor Boris L. Rosovskii, Interdisciplinary Mathematical Sciences, Vol 2, World Scientific (2007), 281–310.

[15] I. Gyöngy, A. Millet, *Rate of Convergence of Space Time Approximations for stochastic evolution equations*, Preprint (2007).

[16] I. Gyöngy, D. Nualart, *Implicit scheme for stochastic parabolic partial differential equations driven by space-time white noise*, Potential Anal. 7, no. 4, 725–757, 1997.

[17] E. Hausenblas, Numerical analysis of semilinear stochastic evolution equations in Banach spaces, J. Comput. Appl. Math. 147, no. 2, 485–516, 2002.

[18] E. Hausenblas, *Approximation for Semilinear Stochastic Evolution Equations*, Potential Analysis, **18** (2003), no 2, 141–186.

[19] E. Hausenblas, *Weak Approximation of Stochastic Partial Differential Equations.* in Capar, U. and Üstünel, A., editor, *Stochastic analysis and related topics VIII. Silivri workshop*, Progress in Probability. Basel: Birkhäuser, 2003.

[20] P.E. Kloeden, E. Platen, *Numerical solution of stochastic differential equations*, Applications of Mathematics (New York), 23. Springer-Verlag, Berlin, 1992.

[21] A. Kohatsu-Higa, *Weak approximations. A Malliavin calculus approach.*, Math. Comp. 70, no. 233, 135–172, 2001.

[22] G. Lord, J. Rougemont, *A Numerical Scheme for Stochastic PDEs with Gevrey Regularity*, IMA J. Num. Anal., **24** (2004), n^o 4, 587–604.

[23] A. Millet, P.L. Morien, *On implicit and explicit discretization schemes for parabolic SPDEs in any dimension*, Stochastic Processes and their Applications **115** (2005), no 7, 1073–1106.

[24] G.N. Milstein *A method with second order accuracy for the integration of stochastic differential equations*, (Russian) Teor. Verojatnost. i Primenen. 23, no. 2, 414–419, 1978.

[25] G.N. Milstein *Weak approximation of solutions of systems of stochastic differential equations.* (Russian) Teor. Veroyatnost. i Primenen. 30 (1985), no. 4, 706–721.

[26] G. N. Milstein, *Numerical integration of stochastic differential equations*, Translated and revised from the 1988 Russian original. Mathematics and its Applications, 313. Kluwer Academic Publishers Group, Dordrecht, 1995.

[27] G. N. Milstein, M. V. Tretyakov, STOCHASTIC NUMERICS FOR MATHEMATICAL PHYSICS, Scientific Computation series, Springer-Verlag, 2004.

[28] D. Nualart, *The Malliavin calculus and related topics*, Springer, 1995.

[29] J. Printems *On the discretization in time of parabolic stochastic partial differential equations*, Math. Model. and Numer. Anal., 35 (6), 1055-1078, 2001.

[30] T. Shardlow, *Numerical methods for stochastic parabolic PDEs*, Numer. Funct. Anal. Optim. **20** (1999), n*o* 1-2, 121–145.

[31] A. Szepessy, R. Tempone, G.Zouraris, *Adaptive weak approximation of stochastic differential equations*, Comm. Pure Appl. Math, 54, 1169-1214, 2001.

[32] D. Talay, *Probabilistic numerical methods for partial differential equations: elements of analysis*, Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995), 148–196, Lecture Notes in Math., **1627**, Springer, Berlin, 1996.

[33] D. Talay, *Discrétisation d'une équation différentielle stochastique et calcul approché d'espérances de fonctionnelles de la solution*, RAIRO Modél. Math. Anal. Numér. 20 (1986), no. 1, 141–179.

[34] Y. Yan, *Galerkin finite element methods for stochastic parabolic partial differential equations*, SIAM J. Numer. Anal. **43** (2005), n^o 4, 1363–1384.

[35] Y. Yan, *Semidiscrete Galerkin approximation for a linear stochastic parabolic partial differential equation driven by an additive noise*, BIT **44** (2004), n^o 4, 829–847.

[36] J.B. Walsh *Finite element methods for parabolic stochastic PDE's*, Potential Anal. **23** (2005), n^o 1, 1–43.